

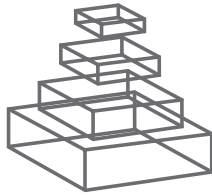
frontiers

RESEARCH TOPICS

STEM CELL GENETIC FIDELITY

Topic Editor
James L. Sherley





frontiers

FRONTIERS COPYRIGHT STATEMENT

© Copyright 2007-2015
Frontiers Media SA.
All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88919-490-2

DOI 10.3389/978-2-88919-490-2

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.

All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

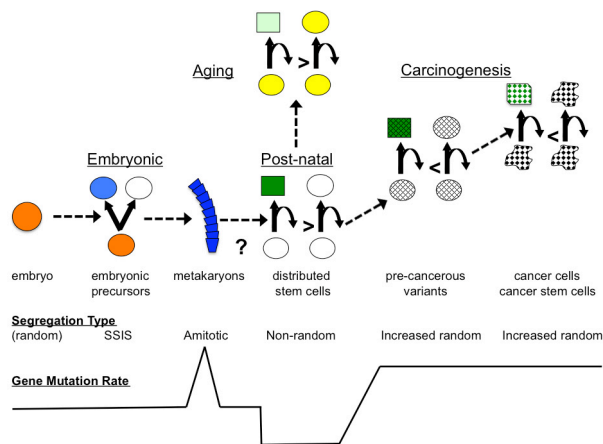
Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

STEM CELL GENETIC FIDELITY

Topic Editor:

James L. Sherley, Asymmetrex, LLC, USA



The mammalian developmental landscape of tissue stem cell genetic fidelity.

The assembled articles (“1-9”) present topics that cover issues of stem cell genetic fidelity throughout mammalian development. In particular, stem cell gene mutation rate is proposed to vary significantly during development. The double-stranded RNA:DNA hybrid genomes of metakaryotic stem cells responsible for early organogenesis are predicted to lead to elevated mutation rates (2). Metakaryotic stem cells, which are the only described morphologically identifiable tissue stem cells, undergo cell divisions without chromatin condensation into chromosomes (i.e., amitotic). The molecular structure of complementary double-strand DNA intrinsically limits such mutation in tissue stem cells (1,2). In contrast, in post-natal tissues in which asymmetric self-renewal by distributed stem cells (i.e., tissue stem cells; 9) predominates over symmetric self-renewal (>), associated non-random sister chromatid segregation is predicted to greatly reduce stem cell mutation rate (4-6). Mutations that disrupt asymmetric self-renewal (<) and non-random segregation (e.g., p53 gene mutations) are predicted to lead to the elevated mutation rates characteristic of cancers (6-8). Mutations occurring in distributed stem cells at the lowest rate are predicted to contribute to tissue aging and cancer (7). Though not directly involved in mutagenesis, a chromosome-specific form of non-random segregation described by the selective sister chromosome segregation (SSIS) hypothesis occurs in embryonic tissue cells (3). SSIS mechanisms may be crucial for epigenetic switches that guide tissue development.

The vision of this Frontiers in Oncology Research Topic on “Stem Cell Genetic Fidelity” had the goal of steeping a diverse range of research perspectives to a first comprehensive synthesis of thought on the questions of how tissue stem cells manage gene mutation rate and the

significance of that management in mammalian evolution and biology, in particular as it relates to tissue cell renewal, carcinogenesis, and aging. The primary focus was determinants of mutation rate in distributed stem cells (DSCs), which encompass all naturally occurring stem cells at all stages of mammalian development. In particular, contributions were sought that considered a broad range of aspects of the immortal DNA strand hypothesis for DSC genetic fidelity. Though proposed in 1975, only in the last decade has this landmark concept in tissue cell biology emerged as a central discussion in DSC research with increasing scrutiny and discussion by an increasing number of laboratories of diverse research perspectives and experimental approaches. With this hypothesis presenting a formidable technical challenge for experimental investigation, as would be expected, both supportive and unsupportive reports have been lining up. In the case of supportive studies, neither the range of applicable tissues nor the responsible molecular mechanisms are known; and the essential genomic process, non-random DNA template strand inheritance by asymmetrically self-renewing DSCs, has been suggested to potentially have other cellular roles besides reducing mutation rate. A major aspiration of this Research Topic was to create the first comprehensive, critical synthesis of current insights and viewpoints on the impact of the immortal DNA strand hypothesis in the history of DSC mutation research. A wide range of article types was considered including historical perspectives, critical reviews, critical commentaries, new hypotheses, new research perspectives, technical advances, and original research reports. Although treatments of the immortal DNA strand hypothesis were the major focus, the desired synthesis required integration of related ideas on mechanisms of DSC mutagenesis and its impact in the evolution of mammals, the emergence of cancers, and stem cell aging. As such, investigators focused on issues in e.g., germ stem cell mutagenesis, effects of environmental mutagens on DSC mutation rate, DSC mutation and tissue aging, determinations of types of mutations in DSCs, and the role of DSC mutation in cancer initiation were invited. Similarly, although the specific goal of the Research Topic was to enlighten DSC genetic fidelity in humans and other mammals, informing contributions based on studies in other model organisms were also welcomed. To achieve even better representation of current experience, advances, and ideas in this field of investigation, these early contributors were encouraged to extend the opportunity to others who shared their interest in advancing our understanding of the mutability of DSCs and its significance in human biology.

Table of Contents

05 Stem Cell Genetic Fidelity

James L. Sherley

Assembled Articles

07 Implications of fidelity difference between the leading and the lagging strand of DNA for the acceleration of evolution

Mitsuru Furusawa

17 Mutator/hypermutable fetal/juvenile metakaryotic stem cells and human colorectal carcinogenesis

Lohith G. Kini, Pablo Herrero-Jimenez, Tushar Kamath, Jayodita Sanghvi, Efren Gutierrez Jr., David Hensle, John Kogel, Rebecca Kusko, Karl Rexer, Ray Kurzweil, Paulo Refinetti, Stephan Morgenthaler, Vera V. Koledova, Elena V. Gostjeva and William G. Thilly

34 Left-right symmetry breaking in mice by left-right dynein may occur via a biased chromatid segregation mechanism, without directly involving the Nodal gene

Stephan Sauer and Amar J. S. Klar

44 Discovering non-random segregation of sister chromatids: the naïve treatment of a premature discovery

Karl G. Lark

50 Comparison of the transcriptomes of long-term label retaining-cells and control cells microdissected from mammary epithelium: an initial study to characterize potential stem/progenitor cells

Ratan K. Choudhary, Robert W. Li, Christina M. Evock-Clover and Anthony V. Capuco

69 Asymmetric cell division and template DNA co-segregation in cancer stem cells

Sharon R. Pine and Wenyu Liu

75 Direct measurements of human colon crypt stem cell niche genetic fidelity: the role of chance in non-Darwinian mutation selection

Haeyoun Kang and Darryl Shibata

85 An APC:WNT counter-current-like mechanism regulates cell division along the human colonic crypt axis: a mechanism that explains how APC mutations induce proliferative abnormalities that drive colon cancer development

Bruce M. Boman and Jeremy Z. Fields

100 Marrow hematopoietic stem cells revisited: they exist in a continuum and are not defined by standard purification approaches; then there are the microvesicles

Peter J. Quesenberry, Laura Goldberg, Jason Aliotta and Mark Dooner



Stem Cell Genetic Fidelity

James L. Sherley*

Asymmetrex, LLC, Boston, MA, USA

*Correspondence: jlsherley@gmail.com

Edited and reviewed by:

Israel Gomy, Institute of Cancer of São Paulo, Brazil

Keywords: tissue stem cell, mutations, non-random segregation, immortal strands

This Frontiers Research Topic, “Stem Cell Genetic Fidelity,” is the product of an attempt to develop a comprehensive integrative treatment of the historical progression of ideas and research advances on the topic of biological mechanisms of gene mutation control in tissue stem cells, with human tissues as a primary interest. In retrospect, this undertaking began with a somewhat ambitious goal for achievement—beginning with its need to recruit authors diverse in both research discipline and generation. In this regard, the scientific characteristics of the final set of contributors are noteworthy. Although many more prospective authors were invited who work in primarily experimental research disciplines relevant to the topic, in the end these are underrepresented in the final volume. The ones for whom this volume’s focus was an effective attractor constitute a select set of cell biologists and molecular biologists whose search for fundamental principles of tissue stem cell biology is driven as much by theoretical modeling approaches as by experimentation.

The deployment of modeling strategies by investigators of the biological principles of genetic fidelity in mammalian tissue stem cells is consistent with the challenges presented by this research. The low tissue fractions of stem cells and the low fractions of the mutations they incur combine to present investigation challenges that defy ideal quantitative analyses. Yet, at the same time, the intrinsic cascades of stem cell-based tissue cell lineages and molecular gene expression hierarchies synergistically amplify minute genetic-cellular deviations in ways that can profoundly influence life and health.

The final cast of contributors is well suited to achieving a particular goal of this Research Topic, which is to bring attention to remarkable advances in cell biology thought that flow from the theoretical genius of four giants in stem cell genetic fidelity research, Lark, Cairns, Potten, and Knudson. Though all applied observation and experimentation for discovering fundamental biological principles, they also employed elegant theoretical modeling as a fine tool for exploring and predicting the workings of tissue cells ahead of experimental approaches and confirmation. Lark, Cairns, and Potten are the most tightly woven into the fabric of ideas on the importance of tissue stem cell genetic fidelity in normal tissue function and carcinogenesis as a result to their shared contributions to the immortal strand hypothesis (Cairns, 1975). Subsequently, Knudson’s two hit hypothesis gave wings to the idea that a key rate limiting factor for the emergence of tumors was two mutations, not only in the alleles of the same gene, but implicitly in the same cell; and most aptly a tissue stem cell (Knudson, 1992). The work of these remarkable

tissue cell biologists—undertaken at the median of stem cell biology history—set the paradigms for current and future research to discover the biological and evolutionary significance of stem cell genetic fidelity mechanisms. In “Stem Cell Genetic Fidelity,” Lark provides a personal account of the key experimental observation that initiated the field of stem cell genetic fidelity research (Lark et al., 1966). As will be noted in reading the Research Topic, several of the contributed articles are descendants of his seminal contribution.

Here, I join Lark with a somewhat personal account of my own that further builds the context for this Frontiers Research Topic. In the fall 1978 as a new student of cancer research, like many before and after, I set myself on a path to “find a cure for cancer.” However, I was more intrigued with mastering the new DNA maps of transforming viral genomes than the changes that they induced in cells. My studies occurred in the last days before the emergence of the concept that cancer was caused by mutations in cellular genes that resembled the oncogenes carried by tumor-forming viruses. Of course, the tenet that gene mutation was an essential aspect of the carcinogenic mechanism was well established long before the first human cancer gene was identified. So, years before I was drawn to cancer research, many cancer scientists were thinking about how carcinogenic mutations arose. However, there was only one who was asking why there were not *more* carcinogenic mutations, and correspondingly a *higher* incidence of human cancers. John Cairns.

In his 1975 report (Cairns, 1975) on the natural history of human cancers, Cairns introduced the fundamental idea of tissue stem cell genetic fidelity and set in motion the essential hypothesis that would trouble the thoughts of cancer scientists and a parallel universe of stem cell biologists for years to come. His proposal that mammalian tissue stem cells must have a unique mechanism to lower their rate of carcinogenic mutation continues to be a disruptive idea in both stem cell biology and cancer biology.

The blade of Cairns’ hypothesis has two sharp edges. First, the “immortal strand hypothesis” cuts through apparent tissue cells and attributes the cell of origin for cancer to tissue stem cells, which are generally physically elusive. Second, the proposed molecular basis for the hypothesis, non-random sister chromatid segregation, slices through geneticists’ essential Mendelian laws of mitotic chromosome segregation. Cairns proposed that *asymmetrically* cycling tissue stem cells ignored Mendel’s previously immutable laws of random assortment and independent segregation. By non-random co-segregation of the complement of sister chromatids with the older template DNA strands, Cairns saw that

asymmetrically cycling tissue stem cells could avoid DNA replication errors, which he predicted to be the most frequent source of carcinogenic mutations.

Cairns' immortal strand hypothesis has met with consternation, skepticism, and outright ridicule. It certainly meets the criteria for a bold, new disruptive idea. Though I had no interest in reading Cairns' paper when it was assigned to my undergraduate Biochemistry course before he arrived as a guest lecturer in the spring of 1978 at Harvard College, in later years I would develop the first theoretical estimate of the extent to which tissue stem cells might lower their mutation rate compared to their differentiating progeny cells (Sherley, 2006); and I would become driven to pursue and promote research to discover the responsible molecular mechanisms. It is my hope that "Stem Cell Genetic Fidelity" will in some measure contribute to this purpose.

REFERENCES

- Cairns, J. (1975). Mutation selection and the natural history of cancer. *Nature* 255, 197–200. doi: 10.1038/255197a0
- Knudson, A. G. (1992). Stem cell regulation, tissue ontogeny, and oncogenic events. *Semin. Cancer Biol.* 3, 99–106.
- Lark, K. G., Consigli, R. A., and Minocha, H. C. (1966). Segregation of sister chromatids in mammalian cells. *Science* 154, 1202–1205. doi: 10.1126/science.154.3753.1202
- Sherley, J. L. (2006). "Mechanisms of genetic fidelity in mammalian adult stem cells," in *Tissue Stem Cells*, eds C. S. Potten, R. B. Clarke, J. Wilson, and A. G. Renehan (New York, NY: Taylor Francis), 37–54.
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 January 2015; accepted: 05 February 2015; published online: 19 February 2015.

Citation: Sherley JL (2015) Stem Cell Genetic Fidelity. *Front. Genet.* 6:51. doi: 10.3389/fgene.2015.00051

This article was submitted to *Cancer Genetics*, a section of the journal *Frontiers in Genetics*.

Copyright © 2015 Sherley. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Implications of fidelity difference between the leading and the lagging strand of DNA for the acceleration of evolution

Mitsuru Furusawa*

Neo-Morgan Laboratory Incorporated, Biotechnology Research Center, Kawasaki, Japan

Edited by:

James L. Sherley, Boston Biomedical Research Institute, USA

Reviewed by:

Matthew Wilkerson, University of North Carolina at Chapel Hill, USA
Amar J. Klar, National Cancer Institute, UK

*Correspondence:

Mitsuru Furusawa, Neo-Morgan Laboratory Incorporated, Biotechnology Research Center, 907 Nogawa, Miyamae-ku, Kawasaki, Kanagawa 216-0001, Japan.
e-mail: furusawa@neo-morgan.com

Without exceptions, genomic DNA of living organisms is replicated using the leading and the lagging strand. In a conventional idea of mutagenesis accompanying DNA replication, mutations are thought to be introduced stochastically and evenly into the two daughter DNAs. Here, however, we hypothesized that the fidelity of the lagging strand is lower than that of the leading strand. Our simulations with a simplified model DNA clearly indicated that, even if mutation rates exceeded the so-called threshold values, an original genotype was guaranteed in the pedigree and, at the same time, the enlargement of diversity was attained with repeated generations. According to our lagging-strand-biased-mutagenesis model, mutator microorganisms were established in which mutations biased to the lagging strand were introduced by deleting the proofreading activity of DNA polymerase. These mutators ("disparity mutators") grew normally and had a quick and extraordinarily high adaptability against very severe circumstances. From the viewpoint of the fidelity difference between the leading and the lagging strand, the basic conditions for the acceleration of evolution are examined. The plausible molecular mechanism for the faster molecular clocks observed in birds and mammals is discussed, with special reference to the accelerated evolution in the past. Possible applications in different fields are also discussed.

Keywords: evolution, acceleration, leading/lagging strand, biased-mutagenesis, replicore, proofreading, polymerase δ

INTRODUCTION

No evolutionary theory has been proposed so far in which the molecular aspect of DNA replication is considered. The main cause of evolution is thought to be mutations accompanying DNA replication. It has been believed that mutations accompanying DNA replication occur stochastically and evenly. Consequently, it can be said that modern evolutionary theories have been built upon this basic thought.

Genomic DNA in all living organisms is replicated by means of the leading/lagging strand system. Therefore, it seems likely that organisms have to pay a higher cost. This is because considerably more enzymes are required for the lagging-strand synthesis than for the synthesis of the leading strand. Theoretically, DNA can be replicated exclusively using the leading strand. Why has nature chosen such a laborious and high cost system? There might be clear evolutionary advantages.

For simplification, a linear DNA with one replication origin (*ori*) at one end, that corresponds to a single replicore in eukaryotes, is used as a model. According to the conventional thought, mutations are introduced stochastically and evenly into the two daughter DNAs. In this situation, if the mutation rate exceeds the so-called threshold value, the population will be extinct before long. Then, this leads to the speculation that organisms might be keeping mutation rates low, and consequently, evolution would advance slowly.

Because of the complexity of the machinery for the lagging-strand synthesis, it can be hypothesized that the fidelity of the

lagging strand might be lower than that of the leading strand. Under this strand-biased-mutagenesis, simulations were carried out. Even if average mutation rates thoroughly exceeded the threshold value, an ancestral genotype was guaranteed in the pedigree and, at the same time, an enlargement of diversity was attained with repeated generations. In other words, the model DNA can produce unchanged progeny maintaining the species, and at the same time, higher mutation rates can provide raw material for evolution (Furusawa and Doi, 1992, 1998). These results support the idea that biased-mutagenesis in the lagging strand may result in the acceleration of evolution without accompanying the extinction of the population (Furusawa and Doi, 1998; Furusawa, 1999).

According to the lagging-strand-biased-mutagenesis model ("disparity model"), "disparity-mutator" microorganisms were provided. A mutator of the intestinal bacterium, *Escherichia coli*, with biased-mutagenesis in the lagging strand was provided, and a disparity mutator of budding yeast, *Saccharomyces cerevisiae*, was also established by deleting the proofreading activity of pol δ . These mutator microorganisms grew normally and showed a quick and extraordinarily high adaptability against very severe circumstances (Tanabe et al., 1999; Shimoda et al., 2006). For instance, an *E. coli* mutator was able to grow colonies even in the presence of saturated concentrations of different kinds of antibiotics (Tanabe et al., 1999).

In this context, the principle of the acceleration of evolution is discussed. The usefulness of a DNA-type genetic algorithm

(GA) with biased-mutagenesis and of disparity mutators of living organisms is also discussed.

ACCELERATION OF EVOLUTION BY DISPARITY-MUTAGENESIS IN DIGITAL ORGANISMS

LEADING/LAGGING-STRANDS REPLICATION SYSTEM PROVIDING EVOLUTIONARY ADVANTAGES

Figure 1 shows a pedigree of a linear chromosomal DNA with a single *ori* at its upper end, i.e., corresponding to a single replicore in higher organisms. It is roughly estimated that several genes, on average, exist in a single replicore. According to our disparity model, biased mutations occur in the lagging strand (Furusawa and Doi, 1992, 1998). The pedigree by the deterministic illustration implies several interesting features; (1) the ancestor has been kept forever; (2) a genotype that once appeared in the past, has been precisely guaranteed in any generation; (3) the threshold of mutation rates is increased (actually, the threshold disappears in this model); (4) even if circumstances changed dramatically, the fittest individual will be selected as a new ancestor and start again to produce a new pedigree as well (**Figure 1**). These outstanding features appear to be beneficial for evolution.

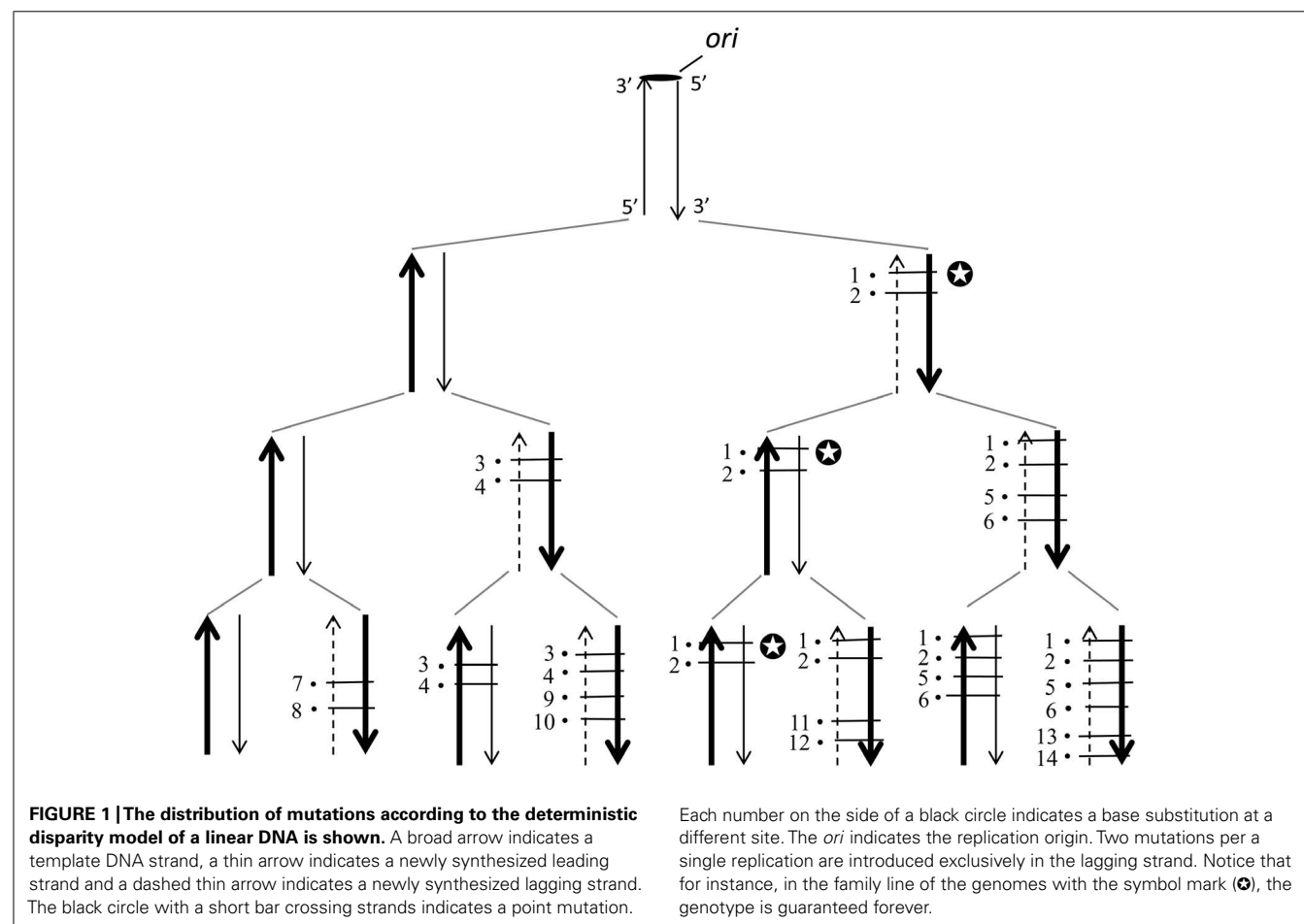
In the conventional model of DNA pedigree, mutagenesis occurs stochastically and evenly in both strands (not shown). Thus, when mutation rates are higher than the threshold value (ca.

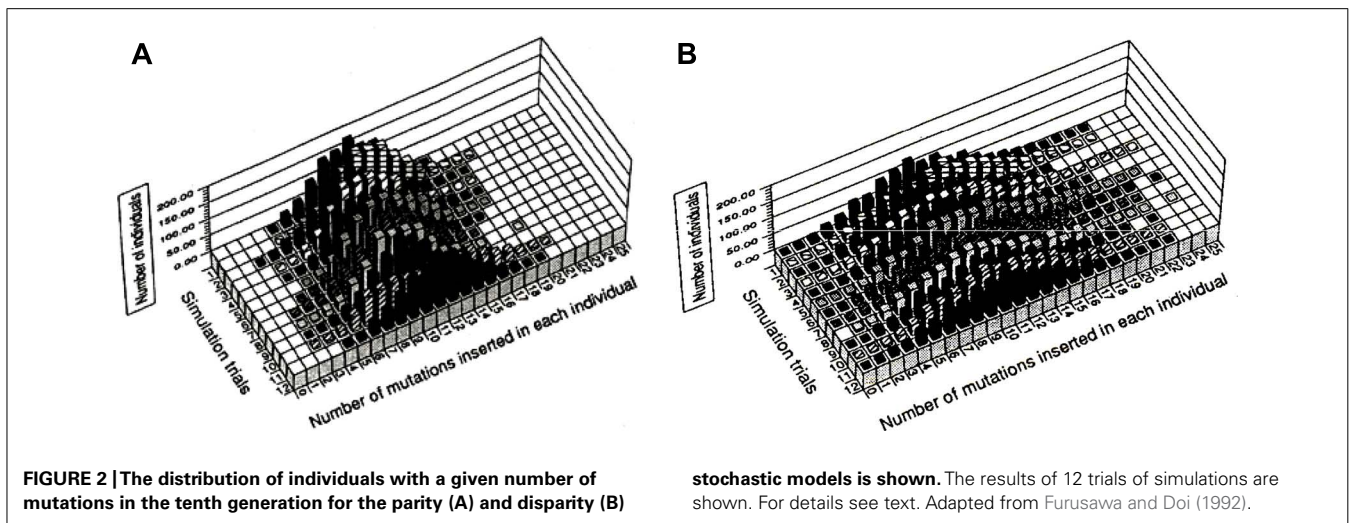
one mutation/daughter DNA/replication), all individuals would eventually die because of the excess accumulation of deleterious mutations.

We compared the distribution of mutations in a parity model (average two mutations/replication being introduced into the leading and the lagging strands) with that of a disparity model (average 1.99 mutations in the lagging strand and 0.01 mutations in the leading strand). For the simulation of this stochastic model, a binominal distribution was used. **Figure 2** summarizes the results of 12 trials at the 10th generation. **Figure 2A** shows the distribution of mutants in the parity model. There is no individual with zero-mutations. In contrast, as shown in **Figure 2B**, the disparity model shows a very flat distribution. The ancestral individuals with zero-mutations are always observable except in the ninth trial and highly mutated mutants comparable with the parity model are produced (Furusawa and Doi, 1992). This flat distribution of mutants including ancestral individuals in the disparity model is expected from the pedigree of the deterministic disparity model shown in **Figure 1**.

DNA-TYPE GENETIC ALGORITHMS WITH BIASED-MUTAGENESIS EFFECTIVELY RESOLVING OPTIMIZATION PROBLEMS

We constructed a GA, named neo-Darwinian algorithm, which consisted of the double-stranded genetic information like DNA

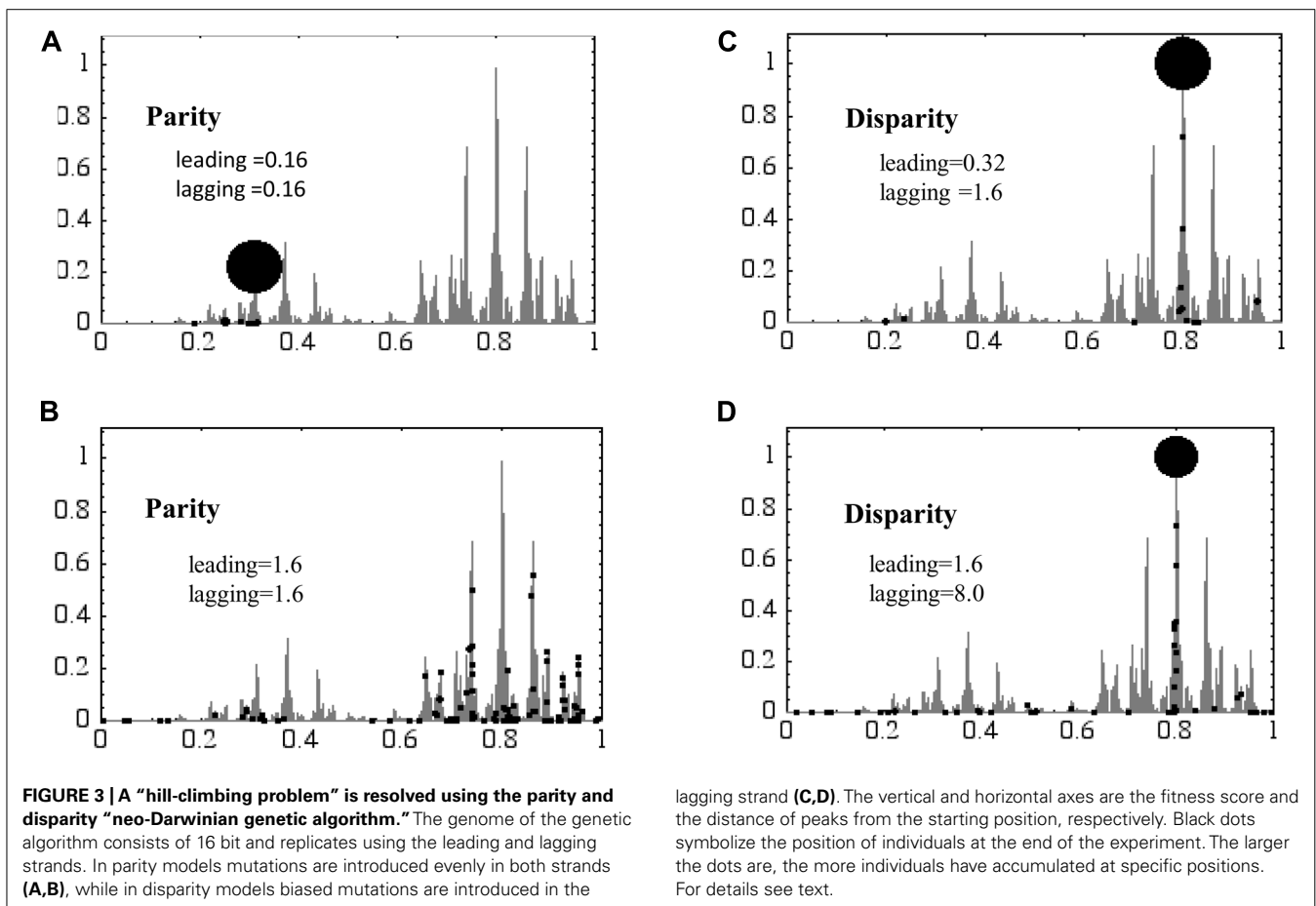




(K. Aoki and M. Furusawa, unpublished). This GA replicates using the leading/lagging strands system. We compared the parity model with the disparity model. A simple “hill-climbing problem” with a rugged landscape was provided. The landscape was made using the gray code which has 686 peaks (Figure 3). An individual that has reached the foot of a given peak can go up the peak by adding

a relatively small number of mutations, while a larger number of mutations are required when an individual moves to another peak at a longer distance.

For simulation, the following conditions were used. The population is always kept at 100, so that 100 individuals having lower fitness scores (FS) have to be omitted from the population after



each replication. Each individual has 16 bit, thus the number of genotypes are $2^{16} = 65,536$. At the beginning of this game, all 100 individuals are located at the left corner, meaning that all of them have the same genotype with an FS of 0.

Figure 3 shows the results at the 100th generation. In the parity model, almost all individuals gather around the top of a small hill of low FS at the 0.16 mutation rates (**Figure 3A**). This is because the mutation rate is too low to look for other higher peaks within 100 rounds of replication. In contrast, when the mutation rate goes up to 1.6, all individuals are scattered all around the landscape, meaning that they cannot keep the positions once obtained because of a too high mutation rate (**Figure 3B**). In other words, in the latter case their genetic information has melted away. The disparity model, in which the mutation rate in the leading strand is 0.32 and lagging strand 1.6, showed very nice adaptability. About 90% of the population has reached the top of the highest hill, and the remaining individuals are located elsewhere (**Figure 3C**). The latter ones are wandering around the peaks searching for higher

peaks that, however, do not exist. When the mutation rate in the leading strand was increased to 1.6 and the lagging strand to 8.0, about 50% of the population has reached the highest hill. The remaining individuals are broadly located in the landscape searching for higher peaks. The results in **Figure 3** clearly showed how the disparity model has a strong adaptability to severe circumstances, especially when mutation rates are high.

Using a similar neo-Darwinian algorithm, a “knap-sack problem” was resolved. The detailed conditions and the results were shown in our previous report (Wada et al., 1993). The digital individual having two chromosomes is sexually reproduced and its DNA has 100 bit. The number of players is 500. The winner is the individual who collects the best combination of “ores” that have the highest value. When the weight of ores packed in the knap-sack exceeds the weight limitation, his boat sinks, acting as selection pressure.

The results are representatively shown in **Figure 4**. At the low mutation rate (0.1), both of the parity and the disparity

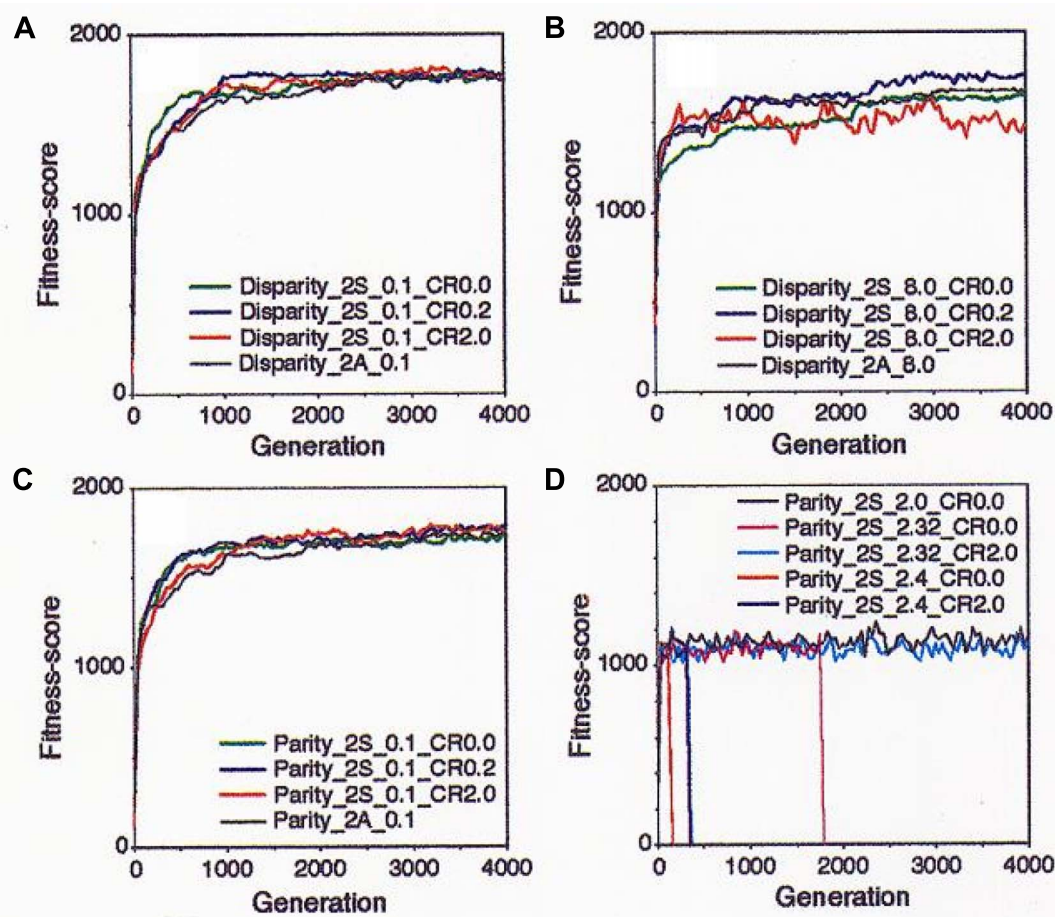


FIGURE 4 | Results of the simulation with the “neo-Darwinian genetic algorithm” in the diploid and sexual world are shown. (A) Disparity individuals with mutation rate (n) = 0.1. Green, no crossover; blue, crossover frequency (CF) = 0.2/chromosome; red, CF = 2.0; and black, asexual and diploid. **(B)** Disparity individuals n = 8.0. Green, no crossover; blue, CF = 0.2; red, CF = 2.0; and black, asexual. **(C)** Parity individuals n = 0.1. Green, no

crossover; blue, CF = 0.2; red, CF = 2.0; and black, asexual. **(D)** Parity individual with various mutation rates. Black, n = 2.0 without crossover; magenta, n = 2.32 without crossover; cyan, n = 2.32 and CF = 2.0; red, n = 2.4 without crossover; and blue, n = 2.4 and CF = 2.0. Adapted from Wada et al. (1993); Copyright 1993, National Academy of Sciences, USA.

individuals showed similar results in that they adapted well and kept high FS as far as tested by 4,000 generations (Figures 4A,C). At higher mutation rates, however, a clear difference was observed between the two models. In the disparity model, even when the mutation rate was as high as 8.0, the individuals showed high adaptability and kept a stable FS. Especially, an appropriate rate of crossover (0.2) increased the FS. In case of the higher crossover rate (2.0), FS was increased with fluctuations. However, the quick increase of the FS curve at early stages of the simulation may mean that the population will quickly occupy the niche, indicating that high crossover rates might act as an advantageous factor for evolution. Individuals with asexual production showed medial FS (Figure 4B). When the parity and the disparity individuals competed with each other in the same niche, the latter quickly drove the former away from the niche (data not shown; see Wada et al., 1993).

On the contrary, the parity model showed a very sensitive reaction to small differences of the mutation rates. At the mutation rate of 2.0, the parity individuals showed a constant FS while the FS value was low compared to that of the disparity. When the mutation rate increased to up to 2.32, the population became extinct by the 2,000th generation. The crossover (0.2) dramatically rescued this extinction. When the mutation rate was 2.4, the players were quickly extinct, but crossover delayed the extinction time (Figure 4D).

The results of the knap-sack problem are summarized as follows: (1) advantageous conditions for the disparity model: small population, strong selection pressure, high mutation rates, diploid sexual reproduction, and competitive world; (2) advantageous conditions for the parity model: large population, weak selection pressure, low mutation rates, haploid asexual reproduction, and non-competitive world. In conclusion, living organisms, especially when environments change dramatically, might decrease the fidelity of the lagging-strand synthesis in order to be able to better adapt to new environments. After the environment has stabilized, the fidelity might be increased up to the usual level. On the other hand, a parity model might be useful for bacteria or yeasts cultured on an agar-plate with sufficient nutrients (Wada et al., 1993).

CO-EXISTENCE OF ERROR-PRONE AND ERROR-LESS POLYMERASES ELIMINATING THE EXISTENCE OF ERROR THRESHOLD FROM EIGEN'S "QUASI-SPECIES"

Eigen's group demonstrated the existence of a critical error threshold using a model RNA (Eigen et al., 1989). Different kinds of RNA polymerase having various fidelities were provided, and each of them was added into a reactor respectively. When the replication reaction reached a stable state, the number of mutations in each RNA molecule was calculated. With increasing error rates of error-prone polymerase, the number of the RNA molecules carrying more mutations naturally increased. Finally, the error rate increased to a critical point just before the genetic information being melted away (the "edge-of-chaos"). If the mutation rate exceeds the critical value, the population immediately falls into a deep chaotic sea, i.e., death (Figure 5A).

However, when a mixture of error-less and error-prone polymerases was used, we obtained completely different results. For

instance, when a mixture of 9% error-less and 91% error-prone polymerase was used, the position of the error threshold was shifted to the right (Figure 5B). When the ratio of error-less polymerase was 10%, the shape of the edge-of-chaos became blunter and the threshold became extinct (Figure 5C). When a mixture of 30% error-less and 70% error-prone polymerases was used the extinction of the population was avoided, even if the average mutation rates thoroughly exceeded the threshold value. Of course, the wild-type with zero-mutations existed forever (Figure 5D; Aoki and Furusawa, 2003). These effects of the mixture of RNA polymerases with different fidelities can be understood corresponding to the effects of biased-mutagenesis in DNA.

Therefore, it can be predicted from the present result that the efficiency of *in vitro* directed evolution of DNA molecules by error-prone polymerase might be considerably increased by adding error-less and error-prone polymerases simultaneously in the PCR.

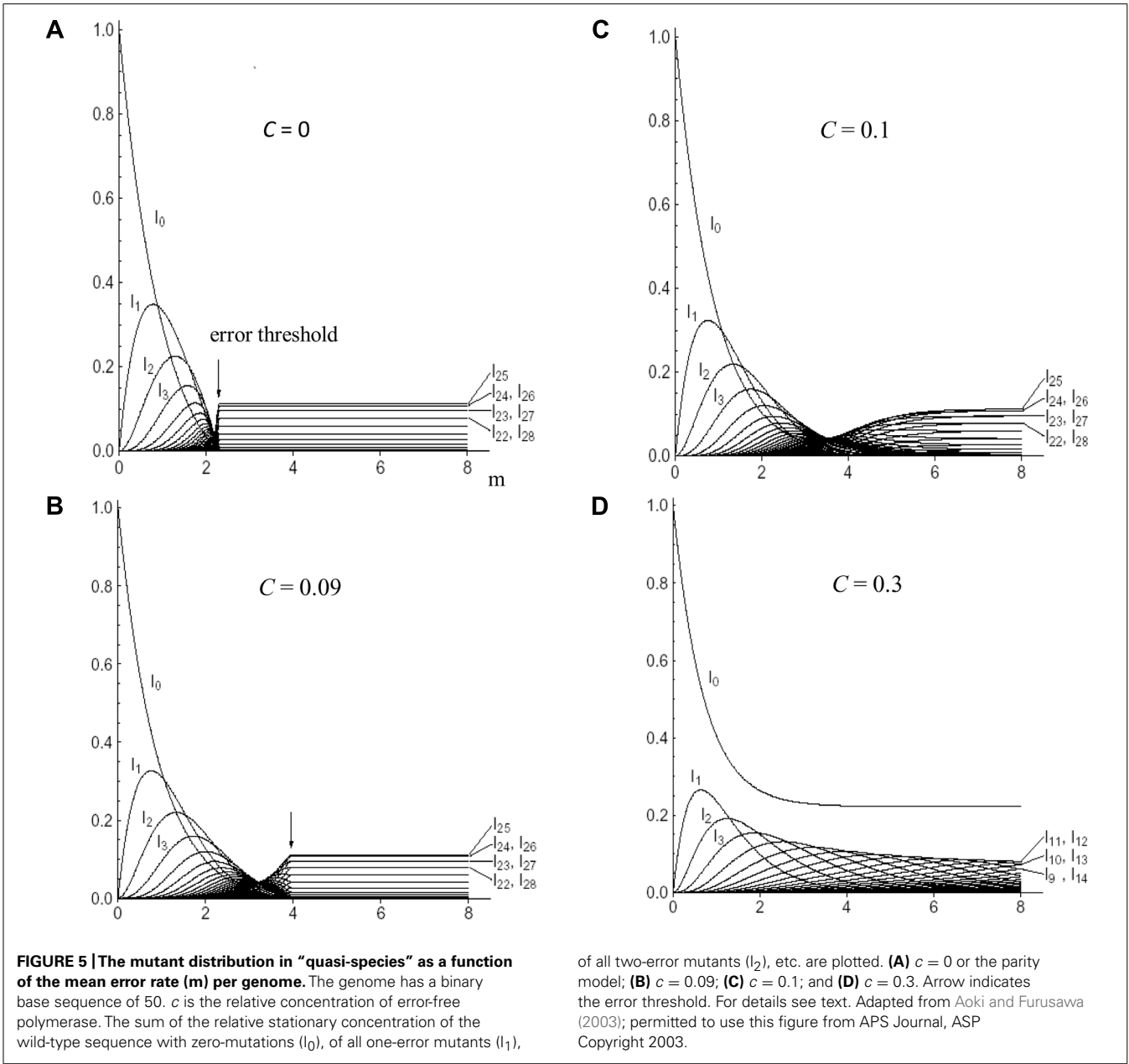
ACCELERATION OF EVOLUTION USING LIVING MICROORGANISMS WITH BIASED-MUTAGENESIS

E. coli DISPARITY-MUTATOR (*dnaQ49*) ACQUIRING AN EXTRAORDINARILY STRONG TOLERANCE TO ANTIBIOTICS

Based on the results of computer simulations, we tried to accelerate evolution using a mutator strain of *E. coli*, *dnaQ49*. The *dnaQ49* strain has a temperature-sensitively mutated gene of *dnaQ* encoded 3'–5' exonuclease (proofreading enzyme). *dnaQ49* shows high mutation rates at 37°C, but nearly normal mutation rates at 24°C. So, we can control mutation rates at will. We showed that mutations were preferentially introduced into lagging strands and that the average mutation rate was estimated to have significantly exceeded the threshold value. Moreover, the relative mutation rate of the lagging strand was about 100 times higher than that of the leading strand (Iwaki et al., 1996). Irrespective of its high mutation rate, *dnaQ49* replicates normally. A normal growth rate of *dnaQ49* is the necessary condition for the experimental acceleration of evolution, since DNA replications are absolutely required for evolution.

For the accumulation of mutations, *dnaQ49* was cultured overnight at 37°C without antibiotics, and on the next day it was cultured on an agar-plate containing different concentrations of antibiotics at 24°C. One colony was picked from the plate which contained the highest concentration of antibiotics, followed by a second cycle of mutagenesis. This cycle of mutagenesis/selection procedure was repeated until no colony was formed. The results are summarized in Table 1.

Most surprisingly, *dnaQ49* was able to survive at the *supersaturation* of different kinds of antibiotics tested; ampicillin, streptomycin, and ofloxacin (a derivative of nalidixic acid). The resultant ampicillin-super-resistant *dnaQ49* that was able to produce colonies at the presence of 30 mg/ml was highly sensitive to other antibiotics, suggesting that the *dnaQ49* was exclusively adapted to the antibiotics used for the respective selection pressure (Table 2). Quinolone is a generic name of a kind of antibiotics including nalidixic acid and ofloxacin, which has anti-gyrase and anti-topoisomerase activity. Quinolone has been clinically used for such a long time that mutant *E. coli* strains have been isolated having acquired tolerance to this drug in patients. Analysis of our



ofloxacin-tolerant *dnaQ49* showed that the sites and the history of point mutations introduced into *gyrA* and *topoIV* genes, which are target genes of quinolone, were coincident with those of the samples from the patients (D83L in *gyrA*, S80R in *topoIV*). Furthermore, no other mutation was found in the limited regions of these two genes as far as sequenced (Tanabe et al., 1999). These results appear to indicate that *E. coli* (*dnaQ49*) evolution is accelerated *in vitro*, and that its results are qualitatively similar to those of naturally occurring evolution.

Excellent adaptability of another *E. coli* mutator in which the lagging-strand-biased-mutagenesis was introduced by mutated *polI* was reported. As *polI* ties Okazaki fragments together, the mutated *polI* introduces mutations exclusively into the lagging strand. When *polI*-mutators having a limited range of mutation

Table 1 Acquired tolerance of <i>dnaQ49</i> to antibiotics.		
	Intact <i>dnaQ49</i> MIC (mg/ml)	Acquired tolerance (mg/ml) of <i>dnaQ49</i>
Ampicillin	0.002	30 ^{a,b}
Streptomycin	0.001	20.06 ^a
Nalidixic acid	0.001	7 ^a
Ofloxacin	0.0000156	3 ^a

This table was made using the data from our previous article (Tanabe et al., 1999). MIC, minimum inhibitory concentration.
^aSaturated concentration.
^bA higher concentration than 30 mg/ml of ampicillin was not available commercially.

Table 2 | Sensitivity of super-ampicillin-tolerant *dnaQ49* to other antibiotics.

	MIC (mg/ml)	
	<i>dnaQ49</i>	Super-ampicillin-tolerant <i>dnaQ49</i>
Ampicillin	0.002	2.048 ^a
Cefotaxime ^b	0.0000313	0.064
Chloramphenicol	0.001	0.0005
Tetracycline	0.00025	0.00025
Rifampicin	0.008	0.002
Streptomycin	0.001	0.0005
Nalidixic acid	0.001	0.0005
Ofloxacin	0.0000156	0.0000156

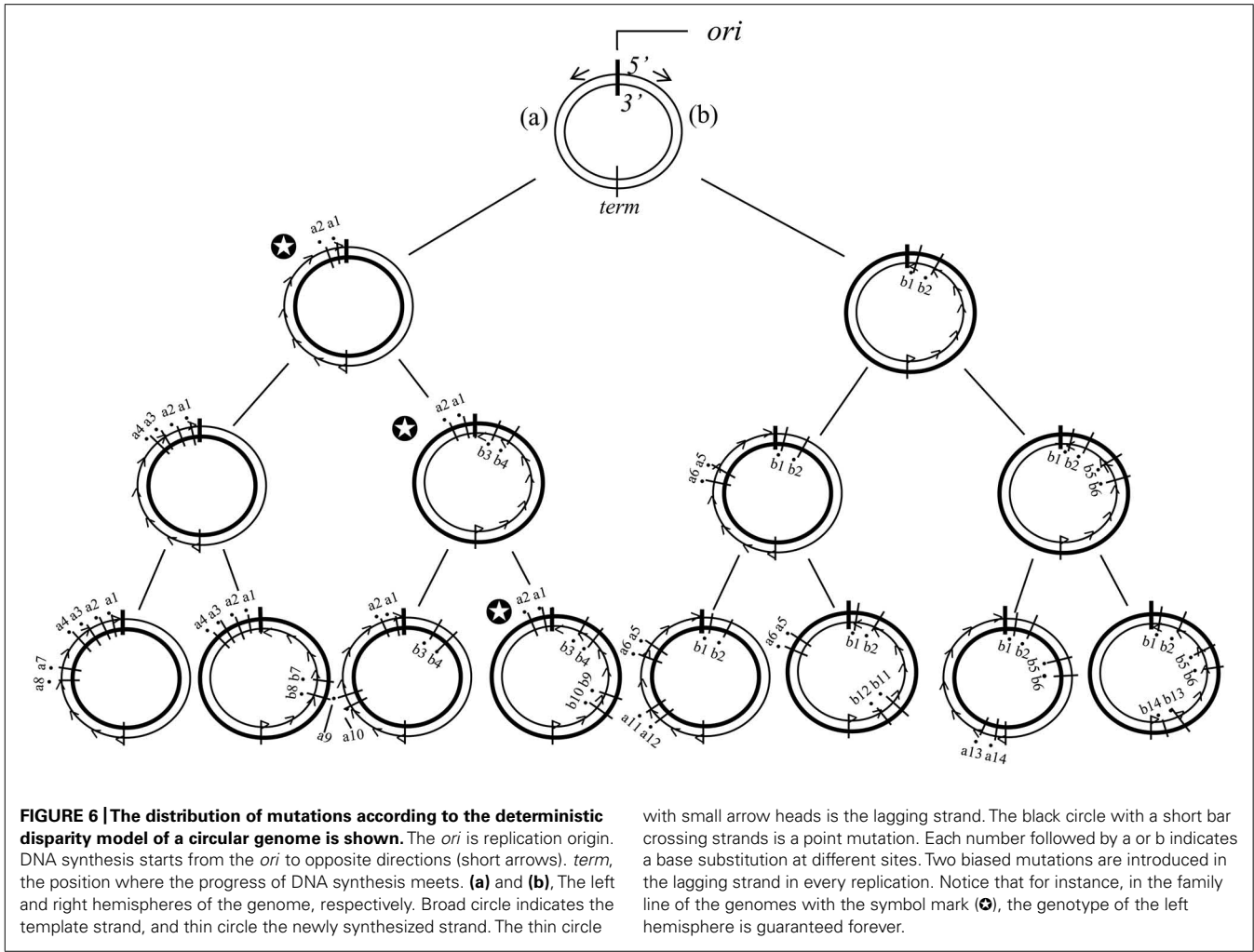
^aGrowing in the presence of 30,000 μg/ml of ampicillin during prolonged culture time.
^bBeta-lactam antibiotic belonging to the group of cephalosporins. Adapted from Tanabe et al. (1999) with minor changes.

rates and having a similar doubling-time were co-cultured, a different mutator strain became the winner of the survival race depending on cases. But, the co-existing wild-type was always driven away from the population (Loh et al., 2010).

Figure 6 shows the deterministic illustrations of the genotypes in the pedigree of the disparity-mutator of bacteria with a single circular genomic DNA. The presence of a given genotype that has once appeared in the left or right hemisphere of the genome is guaranteed forever. Accordingly, at the left or the right hemisphere of the circular genome, the same phenomenon as shown in the pedigree of the disparity model with a liner DNA can be expected (Figure 1). This might explain the reason why the *dnaQ49* mutator showed such a high adaptability against different antibiotics.

ACCELERATION OF EVOLUTION OF YEASTS BY MEANS OF DISPARITY-MUTAGENESIS

Compared to bacteria, eukaryotic cells such as yeasts have a complex genetic constitution, i.e., a plural number of chromosomes and of *oris*, a specific life cycle, meiosis, and recombination, etc. Therefore, firstly, we wondered whether a yeast disparity mutator could behave as *E. coli* mutators do.



For the time being, a disparity mutator of the budding yeast *S. cerevisiae* was used, in which the proofreading activity of *polδ* was deleted by two amino acids substitutions. As lagging strands in the yeast are thought to be exclusively synthesized by *polδ*, this mutator might accumulate mutations in replicores synthesized by mutated *polδ*. This mutant yeast proliferated as well as wild-type. Fortunately, the results were satisfying for us. We quickly isolated adapted mutants that proliferated well at 40°C and survived at 41°C. Then, in order to stop mutagenesis, the mutated *polδ* gene was replaced by the wild-type one by mating with the wild-type yeast. Genetic analyses indicated that at least two genes were concerned with the temperature-resistant phenotype, and we identified one of them, named *hot1* which contributed to the tolerance against 38.5°C (Shimoda et al., 2006).

Recently, Park and colleagues reported that a disparity-mutator of the filamentous yeast *Ashbya gossypii* showed a ninefold improvement in the production of riboflavin compared to the wild-type strain. This mutant was selected by repeated cloning of cells with higher riboflavin-productivity. To establish the mutator, a plasmid-vector harboring the proofreading activity-deleted *polδ* gene was used for transformation. High riboflavin-producing mutants thus obtained were resistant to oxalic acid and hydrogen peroxide as anti-metabolites. Purine and riboflavin biosynthetic pathways were up-regulated, while pathways related to carbon source assimilation, energy generation, and glycolysis were down-regulated. Genes in the riboflavin biosynthetic pathway were significantly over-expressed (Park et al., 2011; Kato and Park, 2012). Importantly, these phenotypes were stable for a significant period of time.

Using disparity mutators of microorganisms, there have been several reports: (1) improvement of the nature of the yeast cell-wall (Abe et al., 2009a), (2) mutated yeasts having multi-stress tolerance which might be useful, e.g., for bio-ethanol production (Abe et al., 2009b), and (3) improvement of the therapeutic glycoprotein productivity of gene-engineered yeasts, which is toxic for the replication of the host yeast (Abe et al., 2009c), and (4) up-regulation of N₂O reductase activity in a nitrogen-fixing bacterium (Itakura et al., 2008). We have also established several tolerant *S. cerevisiae* mutants tolerant against pH 2.5, pH 10.3, 2% of isooctane, and 10% of hexane, toluene, chloroform, cyclohexane, and isooctane, respectively (unpublished).

A STUDY OF MOLECULAR EVOLUTION SUPPORTING OUR DISPARITY MODEL OF EVOLUTION

Katoh et al. (2005) reported that the molecular clock of mammals ran faster than that of other vertebrates. Then, they examined the amino acid substitution rates of *polα*, *polε*, and *polδ*, which mainly contribute to chromosomal DNA replication, obtained from fishes including coelacanth, amphibians, reptiles, birds, and mammals. The result was really exciting for us in that only mammalian *polδ* showed high substitution rates. Moreover, it was recently found out that bird *polδ* also has a higher substitution rate (K. Katoh, personal communication). Thus, it was speculated that the fast run of the bird's and mammalian molecular clocks might be due to the lower fidelity of *polδ*. More surprisingly, amino acid substitutions occurred intensively in the proofreading (exonuclease)-domain of *polδ*. They also showed that the physicochemical nature of

amino acids was changed by those substitutions; strongly suggesting that in the process of bird's and mammalian evolution, the fidelity of *polδ* might be occasionally changed. This change in core amino acids for the proofreading activity would cause an occasional up- and down-regulation of the speed of evolution (Katoh et al., 2005). This unexpected coincidence of our experimental results of acceleration of evolution with those findings of Katoh et al. (2005) appears to indicate that *polδ*, especially its proofreading domain, might play a key role for controlling the speed of evolution.

polδ may also contribute to lagging-strand synthesis in multicellular organisms, though the precise replication mechanism of their genomic DNA is still unclear. In the process of evolution, we can speculate that a similar distribution of mutations shown in Figure 1 might be displayed in each replicore. Because of the highly complex chromosomal constitutions in eukaryotes, we have not yet simulated the effect of biased-mutagenesis on evolution. There is experimental evidence, however, showing that mouse disparity mutators can normally produce descendants without accompanying severe deleterious phenotypes, though cancer predisposition is increased (Albertson et al., 2009; Uchimura et al., 2009). From these facts, we can imagine that organisms might accelerate evolution by means of decreasing the fidelity of the proofreading activity of *polδ*. Or, at least, we can say that evolution might be experimentally accelerated by deleting the proofreading activity of *polδ*.

Judging from the above-mentioned studies of molecular evolution, it is unlikely that evolution is successfully accelerated by artificially decreasing the fidelity of the polymerase domain of *polδ* because most mutations introduced by the polymerase domain of intact *polδ* may not be introduced at random but at the so-called "hot spots." Accordingly, artificial impairment of the polymerase domain by gene-manipulation may mean the disturbance of the natural cause of mutations. There is another evidence that supports the advantage of a disparity mutator for the acceleration of evolution; a disparity mutator of *S. cerevisiae* tends to introduce base-changes by transversion with leading to amino acid substitutions while EMS-mutagenesis tends to introduce transitions mainly from G:C to A:T, not leading to amino acid substitutions (Shiwa et al., 2012). However, we should carefully interpret the result of this report. This is because the yeast mutator harboring mutated *polδ* consists of combined mutations of its polymerase- and proofreading-domains.

After all, according to our data discussed here the best target to be manipulated for attaining experimental acceleration of evolution would be the proofreading (3'-5' exonuclease)-domain of *polδ* gene.

DISCUSSIONS

CONCEPT OF DISPARITY MODEL

The most noticeable feature of our disparity-mutagenesis model would be that parental genotypes are guaranteed by error-less leading strands for realizing more reliable inheritance, while lagging strands make a venture for future evolution by accumulating mutations. The acceleration of evolution might be attained by a simple molecular mechanism, i.e., by amino acid substitutions in the proofreading domain of DNA polymerase contributing to the

lagging-strand synthesis, that lead to a decrease of the fidelity of lagging-strand synthesis. This mechanism might be shared among all organisms.

Our disparity-mutagenesis model would be directly applicable to viruses and prokaryotes which have a single genomic DNA. This is because template strands are always kept in one daughter cell when replicated. In other words, Cairns' "immortal strand model" can basically work, in which the "oldest" strand is always kept in the stem cell (Cairns, 1975; **Figure 6**). As eukaryotes, however, have a plural number of chromosomes in each cell and a plural of *oris* on a single chromosome, situations must be much more complex. The effect of disparity-mutagenesis on eukaryotic evolution remains to be examined.

EFFECT OF CROSSOVER ON DISPARITY-MUTAGENESIS MODEL

Another point to be considered is that of crossover in sexually reproducing organisms. A limited number of crossovers between each pair of homologous chromosomes occur once during the maturation of germ cells. As shown in **Figure 4**, we can expect appropriate rates of crossover occurring inside a replicore to increase FS. However, usually the crossover might decrease the FS. Anyway, the effect of crossover on FS remains to be examined.

HIGH MUTATION RATES IN HUMAN AND DISPARITY MODEL

It has been reported that the spontaneous mutation rate of human is inconceivably high, more than 100 mutations/cell/generation. There might be at least 2~3 deleterious mutations which certainly exceed the error threshold (Kondrashov, 1995; Crow, 1997; Drake et al., 1998; Eyre-Walker and Keightley, 1999; Kong et al., 2012). Why have we not died? These excess mutations are believed to be mainly introduced during spermatogenesis (Kong et al., 2012). If we hypothesize that these large numbers of mutations are introduced evenly into the chromosomes of infants, the reason why we are still here would be hard to be explained. Therefore, DNA polymerases that act during spermatogenesis would remain to be analyzed from the viewpoint of disparity-mutagenesis.

REPLICORE, DISPARITY-MUTAGENESIS, AND EVOLUTION

According to our disparity model of evolution, we can suppose the situation in that a given advantageous replicore once appeared is basically guaranteed in the population for a long time even when mutation rates are high. As if an allele drifts in the population, the replicore as a "stable" genetic unit might be drifting in the population and would be used, by chance, by someone and somewhere. A single replicore usually consists of several genes. Moreover, genes with similar or related functions tend to locate at positions close to each other and/or at a single replicore, e.g., the IgG heavy chain (Brown et al., 1978), β -globin family (Maniatis et al., 1982), *Drosophila* Antennapedia complex (Harding et al., 1985), mouse Hox (Graham et al., 1989), ribosomal RNA (Long and David, 1980), tRNA (Clarkson, 1980), histones genes (Hentschel and Birnstiel, 1981), etc. Thus, the drift of advantageous replicores might offer more drastic effects on evolution than a single gene. Another important condition for guaranteeing the existence of stable replicores in the population would be that germ-lines must have been keeping the positions of *oris* unchanged

for a long time. It seems likely that the existence of the replicore as a stable genetic unit would serve as a main cause of "linkage disequilibrium."

APPLICATIONS OF DISPARITY MODEL

DNA-type GA as presented here would be useful for resolving optimization problems with environments changing over time, and the disparity-mutator organisms would be useful not only for the improvement of organisms but also for the improvement of bio-products in medical and industrial fields.

IMPLICATIONS OF THE ACCELERATION OF EVOLUTION

The concept of artificial acceleration of evolution presented here would serve as a possible tool for opening the door of the black-box existing between gene and organism. Because, we can now see better the process of genomic and phenotypic changes within a living organism (Furusawa, 1999).

NOVEL BIOLOGICAL FUNCTIONS OF DOUBLE-STRANDED DNA STRUCTURE

Last of all, Cairns' "immortal strand hypothesis" for cancer prevention (Cairns, 1975), Klar's "somatic strand-specific imprinting and selective chromatid segregation (SSIS) model" for the determination of differentiation (Klar, 2007), and the present disparity-mutagenesis model would share a common concept. There exists a basic thread common to all of them, i.e., these three lines of research have tried to seek for additional biological implications of the double-stranded structure of the DNA molecule (Furusawa, 2011). Their new paradigm seems to involve implications far beyond conventional molecular biology.

CONCLUSION

- (1) Living organisms are able to represent two contradictory phases; an unchangeable (heredity) and a changeable phase (evolution). According to our disparity theory of evolution, the ultimate cause of the precise heredity would be traced in the leading strand of high fidelity, and that of evolution in the lagging strand of low fidelity.
- (2) A double-stranded DNA and its replication machinery can be regarded as a sort of GA. Thus, the fidelity difference between these two strands would act as a factor which increases the resolution ability for optimization problems in environments.
- (3) The disparity-mutagenesis model can predict that the speed of evolution appears to be regulated by mutations introduced in the proofreading domain of pol δ during the evolutionary process.
- (4) Artificially decreasing the proofreading activity of pol δ , which may cause increasing the rate of biased-mutagenesis in the lagging strand, would give rise to the experimental acceleration of evolution in eukaryotes.

ACKNOWLEDGMENTS

The author thanks Dr F. Rueker for his critical reading of the manuscript and for useful suggestions. The author also thanks Drs A. Uchimura and T. Yagi for their fruitful discussions about high mutation rates in human.

REFERENCES

- Abe, H., Fujita, Y., Chiba, Y., Jigami, Y., and Nakayama, K. (2009a). Upregulation of genes involved in gluconeogenesis and the glyoxylate cycle suppressed the drug sensitivity of an N-glycan-deficient *Saccharomyces cerevisiae* mutant. *Biosci. Biotechnol. Biochem.* 73, 1398–1403.
- Abe, H., Fujita, Y., Takaoka, Y., Kurita, E., Yano, S., Tanaka, N., et al. (2009b). Ethanol-tolerant *Saccharomyces cerevisiae* strains isolated under selective conditions by over-expression of a proofreading-deficient DNA polymerase delta. *J. Biosci. Bioeng.* 108, 199–204.
- Abe, H., Takaoka, Y., Chiba, Y., Sato, N., Ohgiya, S., Itadani, A., et al. (2009c). Development of valuable yeast strains using a novel mutagenesis technique for the effective production of therapeutic glycoproteins. *Glycobiology* 19, 428–436.
- Albertson, T., Ogawa, M., Bugni, J., Hays, L., Chen, Y., Wang, Y., et al. (2009). DNA polymerase ϵ and δ proofreading suppress discrete mutator and cancer phenotypes in mice. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17101–17104.
- Aoki, K., and Furusawa, M. (2003). Increase in error threshold for quasispecies by heterogeneous replication accuracy. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 68(Pt 1), 031904.
- Brown, E., Iqbal, M., Stuart, S., Hatton, K., Valinsky, J., and Schildkraut, C. (1978). Rate of replication of the murine immunoglobulin heavy-chain locus: evidence that the region is part of a single replicon. *Mol. Cell. Biol.* 7, 450–457.
- Cairns, J. (1975). Mutation selection and the natural history of cancer. *Nature* 255, 197–200.
- Clarkson, S. (1980). "Their structure, activity and regulation," in *Eukaryotic Genes*, eds N. Maclean and S. P. Gregory (London: Butterworth), 239–266.
- Crow, J. (1997). The high spontaneous mutation rate: is it a health risk? *Proc. Natl. Acad. Sci. U.S.A.* 94, 8380–8386.
- Drake, J. W., Charlesworth, B., Charlesworth, D., and Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics* 148, 1667–1686.
- Eigen, M., McCaskill, J., and Schuster, P. (1989). The molecular quasispecies. *Adv. Chem. Phys.* 75, 149–263.
- Eyre-Walker, A., and Keightley, P. (1999). High genomic deleterious mutation rates in hominids. *Nature* 397, 344–347.
- Furusawa, M. (1999). *DNA's Exquisite Evolutionary Strategy*. Tokyo: Kodansha.
- Furusawa, M. (2011). Implications of double-stranded DNA structure for development, cancer and evolution. *OJGen* 1, 78–87.
- Furusawa, M., and Doi, H. (1992). Promotion of evolution: disparity in the frequency of strand-specific misreading between the lagging and leading strands enhances disproportionate accumulation of mutations. *J. Theor. Biol.* 157, 127–133.
- Furusawa, M., and Doi, H. (1998). Asymmetrical DNA replication promotes evolution: disparity theory of evolution. *Genetica* 102–103, 333–347.
- Graham, A., Papalopulu, N., and Krumlauf, R. (1989). The murine and *Drosophila* homeobox gene complexes have common features of organization and expression *Cell* 57, 367–378.
- Harding, K., Wedeen, W., and Levine, M. (1985). Spatially regulated expression of homeotic genes in *Drosophila*. *Science* 20, 1236–1242.
- Hentschel, C., and Birnstiel, M. (1981). The organization and expression of histone gene families. *Cell* 25, 301–313.
- Itakura, M., Tabata, K., Eda, S., Mitsui, H., Murakami, K., Yasuda, J., et al. (2008). Generation of *Bradyrhizobium japonicum* mutants by selection after introduction of mutated *dnaQ* gene. *Appl. Environ. Microbiol.* 74, 7258–7264.
- Iwaki, T., Kawamura, A., Ishino, Y., Kohno, K., Kano, Y., Goshima, N., et al. (1996). Preferential replication-dependent mutagenesis in the lagging DNA strand in *Escherichia coli*. *Mol. Gen. Genet.* 251, 657–664.
- Kato, T., and Park, E. (2012). Riboflavin production by *Ashbya gossypii*. *Biotechnol. Lett.* 34, 611–618.
- Katoh, K., Kuma, K., Iwabe, N., and Miyata, T. (2005). "Putative decline of DNA polymerase δ in mammalian lineage," in *Proceedings of the 28th Conference on The Molecular Biology Society of Japan*, December 7–10, 2005, Fukuoka, 2P-0257.
- Klar, A. (2007). Lessons learned from studies of fission yeast mating-type switching and silencing. *Annu. Rev. Genet.* 41, 213–236.
- Kondrashov, S. (1995). Contamination of the genomes by very slightly deleterious mutations. Why have we not died 100 times over? *J. Theor. Biol.* 175, 583–594.
- Kong, A., Frigge, M., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., et al. (2012). Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* 488, 471–475.
- Loh, E., Salk, J., and Loeb, L. (2010). Optimization of DNA polymerase mutation rates during bacterial evolution. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1154–1159.
- Long, E., and David, I. (1980). Repeated genes in eukaryotes. *Annu. Rev. Biochem.* 49, 727–764.
- Maniatis, T., Melon, P., Parker, V., Proudfoot, N., and Seed, B. (1982). "Molecular genetics of human globin gene expression," in *Molecular Genetics of Neuroscience*, eds F. O. Schmitt, S. J. Bird, and F. E. Bloom (New York: Raven Press), 87–101.
- Park, E., Ito, Y., Nariyama, M., Sugimoto, T., Lies, D., and Kato, T. (2011). The improvement of riboflavin production in *Ashbya gossypii* via disparity mutagenesis and DNA microarray analysis. *Appl. Microbiol. Biotechnol.* 91, 1315–1326.
- Shimoda, C., Itadani, A., Sugino, A., and Furusawa, M. (2006). Isolation of thermotolerant mutants by using proofreading-deficient DNA polymerase δ as an effective mutator in *Saccharomyces cerevisiae*. *Genes Genet. Syst.* 81, 391–397.
- Shiwa, Y., Fukushima-Tanaka, S., Kasahara, K., Horiuchi, T., and Yoshikawa, H. (2012). Whole-genome profiling of a novel mutagenesis technique using proofreading-deficient DNA polymerase δ . *Int. J. Evol. Biol.* 2012, 860797.
- Tanabe, K., Kondo, T., Onodera, Y., and Furusawa, M. (1999). A conspicuous adaptability to antibiotics in the *Escherichia coli* mutator strain, *dnaQ49*. *FEMS Microbiol. Lett.* 176, 191–196.
- Uchimura, A., Hidaka, Y., Hirabayashi, T., Hirabayashi, M., and Yagi, T. (2009). DNA polymerase δ is required for early mammalian embryogenesis. *PLoS ONE* 4, e4184. doi: 10.1371/journal.pone.0004184
- Wada, K., Doi, H., Tanaka, S., Wada, Y., and Furusawa, M. (1993). A neo-Darwinian algorithm: asymmetrical mutations due to semiconservative DNA-type replication promote evolution. *Proc. Natl. Acad. Sci. U.S.A.* 90, 11934–11938.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 August 2012; accepted: 27 September 2012; published online: 16 October 2012.

Citation: Furusawa M (2012) Implications of fidelity difference between the leading and the lagging strand of DNA for the acceleration of evolution. *Front. Oncol.* 2:144. doi: 10.3389/fonc.2012.00144

This article was submitted to *Frontiers in Cancer Genetics*, a specialty of *Frontiers in Oncology*.

Copyright © 2012 Furusawa. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



Mutator/hypermutable fetal/juvenile metakaryotic stem cells and human colorectal carcinogenesis

Lohith G. Kini¹, Pablo Herrero-Jimenez², Tushar Kamath¹, Jayodita Sanghvi³, Efen Gutierrez Jr.⁴, David Hensle⁵, John Kogel⁵, Rebecca Kusko⁶, Karl Rexer⁷, Ray Kurzweil⁸, Paulo Refinetti⁹, Stephan Morgenthaler⁹, Vera V. Koledova¹, Elena V. Gostjeva¹ and William G. Thilly^{1*}

¹ Laboratory for Metakaryotic Biology, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA

² SLC, Oakville, ON, Canada

³ Department of Bioengineering, Stanford University, Stanford, CA, USA

⁴ Massachusetts General Hospital for Children, Boston, MA, USA

⁵ Private Ind., Chicago, IL, USA

⁶ Boston University Medical Center, Boston, MA, USA

⁷ Rexer Analytics, Winchester, MA, USA

⁸ Kurzweil Technologies, Inc., Wellesley, MA, USA

⁹ Department of Mathematics, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Edited by:

James L. Sherley, Boston Biomedical Research Institute, USA

Reviewed by:

Bruce M. Boman, University of Delaware, USA

James DeGregori, University of Colorado School of Medicine, USA

*Correspondence:

William G. Thilly, Laboratory in Metakaryotic Biology, Department of Biological Engineering, Massachusetts Institute of Technology, Bldg. 16-771, 77 Massachusetts Avenue, Cambridge, MA 01890, USA
e-mail: thilly@mit.edu

Adult age-specific colorectal cancer incidence rates increase exponentially from maturity, reach a maximum, then decline in extreme old age. Armitage and Doll (1) postulated that the exponential increase resulted from “*n*” mutations occurring throughout adult life in normal “cells at risk” that initiated the growth of a preneoplastic colony in which subsequent “*m*” mutations promoted one of the preneoplastic “cells at risk” to form a lethal neoplasia. We have reported cytologic evidence that these “cells at risk” are fetal/juvenile organogenic, then preneoplastic metakaryotic stem cells. Metakaryotic cells display stem-like behaviors of both symmetric and asymmetric nuclear divisions and peculiarities such as bell shaped nuclei and amitotic nuclear fission that distinguish them from embryonic, eukaryotic stem cells. Analyses of mutant colony sizes and numbers in adult lung epithelia supported the inferences that the metakaryotic organogenic stem cells are constitutively mutator/hypermutable and that their contributions to cancer initiation are limited to the fetal/juvenile period. We have amended the two-stage model of Armitage and Doll and incorporated these several inferences in a computer program CancerFit v.5.0. We compared the expectations of the amended model to adult (15–104 years) age-specific colon cancer rates for European-American males born 1890–99 and observed remarkable concordance. When estimates of normal colonic fetal/juvenile APC and OAT gene mutation rates ($\sim 2\text{--}5 \times 10^{-5}$ per stem cell doubling) and preneoplastic colonic gene loss rates ($\sim 8 \times 10^{-3}$) were applied, the model was in accordance only for the values of $n = 2$ and $m = 4$ or 5.

Keywords: stem cells, metakaryotic, mutator, hypermutable, cancer model

INTRODUCTION

The age-specific incidence of colorectal cancer is here reconsidered in terms of stem cell growth and oncomutation in normal stem cells of the fetal/juvenile period (tumor initiation) and thereafter in the preneoplastic stem cells of colonic adenomas (tumor promotion) from which lethal colonic adenocarcinomas arise. This reconsideration was prompted by discoveries concerning metakaryotic stem cells in organogenesis and carcinogenesis (2–5) and by evidence that normal organogenic stem cells experience

unexpectedly high rates of genetic changes required for tumor initiation (6).

OVERVIEW OF ORGANOGENESIS AND CARCINOGENESIS

The human body arises from a single fertilized egg and in a series of cell divisions creates a body mass of some 2^{44} cells at maturity. However, these cells are not homogeneous but are apportioned among the organs each containing several distinct tissue layers containing in turn one or more histologically recognizable cell types. The epithelial layers of many solid organs, such as those of the gastrointestinal tract, the lung, breast and prostate are of special interest because it appears that the vast majority of lethal tumors arise from these layers. Each of these epitheloid or adenocarcinomatous (gland-like) tumors displays histological organization that, however distorted, resembles the histological organization of that organ during fetal growth and development (7). In the case of the human colon, histologic alterations at multiple foci in

Abbreviations: APC, adenomatous polyposis coli gene; FAPC, familial adenomatous polyposis coli; OAT, sialylmucin acid O-acetyl transferase gene; *t*, age interval of death; *h*, cohort (birth years, gender, ethnic group); *y*, calendar year of death; MOR(*h*,*t*), number of deaths; POP(*h*,*t*), number in population; OBS(*h*,*t*), raw death rate, MOR(*h*,*t*)/POP(*h*,*t*); OBS*(*h*,*t*), raw death rate corrected for coincident deaths, corrected death rate; INC(*h*,*t*), corrected death rate adjusted for survival of disease to approximate incidence; CAL(*h*,*t*), values of incidence predicted by amended model; GOF(*h*,*t*), goodness of fit of CAL(*h*,*t*) to INC(*h*,*t*).

the embryonic mid- and hind gut begin at about the 5th week of gestation; by the 25th week one may observe the parallel array of crypts that open to the colonic lumen numbering about one million ($\sim 2^{20}$) in a newborn and about ten million ($\sim 2^{23,25}$) at maturity with each crypt containing about two thousand ($\sim 2^{11}$) epithelial cells.

In adult epithelial tissues slowly growing epitheloid colonies (adenomas) are observed from which more rapidly growing colonies emerge (adenocarcinomas). From some, but not all, adenocarcinomas even more rapidly growing metastases emerge that are distributed throughout the body. It appears that each succeeding step represents an event in a single stem cell such that each form of lesion is clonal.

METAKARYOTIC STEM CELLS OF ORGANOGENESIS AND CARCINOGENESIS

Embryonic stem cell lines derived from human blastulas are eukaryotic cells. They have spheroidal nuclei enclosed in a nuclear membrane that exists as an organelle within the cell cytoplasm. Their DNA content doubles in an S-phase that is completed a few hours before the condensation of chromosomes in prophase marks the beginning of mitosis by which the two sister genomes are segregated during cell fission to form two sister cells.

But beginning in the fourth- to fifth-week of gestation and in all preneoplasias and neoplasias examined a different kind of cell is found in which the genomic DNA is contained in a large, hollow, bell shaped structure. These appear to arise from precursor cells with spherical nuclei, i.e., resembling eukaryotic embryonic stem cells, in which a belt of condensed chromatin marks the beginning of an amitotic process in which two separate facing hemispheres are created each containing the diploid amount of DNA. Soon these bell shaped nuclei are enclosed in a sarcomeric pod or tubular syncytium (3, 4). In the ~ 5 th–12th week of gestation the number of bell shaped nuclei per tubular syncytium increases by a form of symmetric amitosis resembling one paper cup separating from another. The number of syncytia increases rapidly and distribute non-randomly in space and time within fetal meta-organs up to about the 12th week. Then the syncytia disappear but the bell shaped nuclei persist as mononuclear cells with oblate spheroidal mucinous cytoplasms. These continue to increase by symmetric cup-from-cup amitoses without any detectable condensation of chromosomes. Bell shaped nuclei are appended to, rather than enclosed in, the cytoplasmic organelles. DNA doubles not in an S-phase preceding nuclear fission as in eukaryotes but during and after amitotic segregation of bell shaped sister nuclei (2–4). The genomes of bell shaped nuclei between doublings appear to be organized in a set of circular structures each containing one or more homologously paired chromosomal elements specifically end-joined at their telomeres (5).

Both in the syncytial and post syncytial period of development bell shaped nuclei of the developing colon undergo asymmetric amitoses in which any of at least nine distinct morphologic forms of closed nuclei emerge from bell shaped nuclei. These subsequently undergo a series of DNA doublings and mitoses that create the eukaryotic cells that populate the crypts of the colonic epithelium throughout the second and third fetal trimesters and the juvenile period. Cells with bell shaped nuclei external to

cytoplasmic organelles are located at the bases of colonic crypts both in the developing fetus and colonic adenocarcinomas. After maturity cells with bell shaped nuclei are rarely found at the base of colonic crypts or epithelia of any organs observed to date but are invariably found in incipient micro-colonies in colonic adenomas and at the base of crypts in colonic adenocarcinomas and their derived metastases (3).

Insofar as these cells displayed cytologic behaviors distinguishing them from eukaryotic cells, they were designated as *metakaryotic* cells because they were first observed in the formation of meta-organs during the 5th- to 12th-week of gestation (3). Their discovery in the development of many organs including those of the developing digestive and nervous systems and derived tumors as well as their appearance in a developing plant marked them as an important and general part of organogenesis and carcinogenesis with evolutionary origins preceding the separation of plants and animals (4).

The demonstration that the bell shaped nuclei of metakaryotic cells underwent both symmetric amitoses to increase in number and asymmetric amitoses to form the mitotic parenchymal cells of the organ or tumor marked them as organogenic or carcinogenic stem cells (2–4). These observations confirmed and extended the observations and interpretations of Child (1907) that the gonadal development of a sheep tapeworm involved a series of amitotic divisions followed by a mitotic expansion and directly contradicted the opinion of Boveri (8) that tissue development and tumor growth depended solely on mitotic fission.

HIGH RATES OF STEM CELL MUTATION IN A DEVELOPING HUMAN ORGANS

Measurement of five specific nuclear point mutations and 17 specific mitochondrial point mutations in the DNA of micro-anatomical samples of adult human lung epithelia revealed unexpectedly high numbers of somatic mutations in the *TP53*, *KRAS*, and *HPRT1* nuclear genes and in the mitochondrial genome (bp10,100–10,101) (6, 9). Plotting mutant colony numbers as a function of colony size for the lung data revealed a Luria–Delbruck distribution (10). This finding indicated that the mutations occurred at a constant rate per organogenic stem cell doubling during the exponential growth of the organ, i.e., the fetal/juvenile period. Furthermore, mutant fractions in the lung epithelium did not increase with age indicating that mutation rates in maintenance stem cells of lung turnover units were much lower than in organogenic stem cells. Consistent with this interpretation was the observation that overall lung mutant colony numbers of any size did not significantly increase with age in adults. Using these data we estimated a gene-inactivation mutation rate per lung stem cell doubling of about $2\text{--}4 \times 10^{-4}$ for the *TP53* and *HPRT1* genes.

Histological enumeration of colonic polyps carrying a somatic mutation of the *APC* gene in FAP patients' colons (11) and of colonic crypts carrying a somatic mutation in the sialomucin acid *O*-acetyl transferase (*OAT*) gene (12) provided data that have allowed us to estimate that gene-inactivation mutation rates for stem cells of the developing colon are about $2\text{--}5 \times 10^{-5}$. As the gene-inactivation rate of human B-cells grown in culture has been estimated as $\sim 2 \times 10^{-7}$ per cell doubling (13) it is clear

that fetal/juvenile stem cell mutation rates are some 100 (colon) to 1000× (lung) higher than previously associated with human eukaryotic cells. Levels of loss of heterozygosity for polymorphic markers in colonic preneoplasia of ~0.25 allow estimation of preneoplastic gene deletion rates of about 8×10^{-3} per gene copy per preneoplastic stem cell doubling (14). Based on these observations we inferred that high rates of genetic change or “genomic instability,” generally associated with preneoplastic and neoplastic growths, are a characteristic of the metakaryotic stem cells of human organogenesis.

MATERIALS AND METHODS

BIOLOGICALLY BASED ALGEBRAIC CANCER MODELS

We have incorporated these new biological findings/inferences in a revised model of carcinogenesis adopting and extending the two-stage model first proposed and expressed in algebraic form by Armitage and Doll (1): “*n*” *initiating* mutations in an organogenic metakaryotic stem cell permit it to (a) avoid metamorphosis to a non-growing maintenance stem cell at maturity and (b) continue to grow at approximately the juvenile growth rate as a preneoplastic colony; “*m*” *promoting* events in a preneoplastic metakaryotic stem cell permit it to form a rapidly growing lethal neoplastic colony.

The model so revised permits, for the first time, the use of key parameters derived from clinical observations: the number of stem cell doublings during organogenesis (a_{\max}) and the rates per metakaryotic organogenic stem cell doubling of *n* events necessary for initiation (R_i, R_j, \dots, R_n) and the rates per preneoplastic metakaryotic stem cell doubling of *m* events necessary for promotion (R_A, R_B, \dots, R_m). Comparison of quantitative predictions based on the revised model to the age-specific incidence rate for that cancer in a population in which data are available for the entire adult life span of ~15–104 years tests the general accuracy of the assumptions incorporated in the model.

One-stage models

Cohnheim inferred from similarity of histologic organization between second trimester fetuses and adenocarcinomas that tumors originated in cells responsible for embryonic/fetal growth and differentiation (7). Boveri inferred from the presence of altered chromosome numbers and structures in some primary tumors and metastases that tumors involved chromosomal/genetic changes (8). In the early 1950s several theorists sought to reconcile the increasing rate of cancers with adult age in terms of new understandings of genetics and genetic change particularly the demonstration that human tumors were of clonal origin. They posited that any single cell at risk in a static, adult cell population could be created from normal “cells at risk” by “*n*” required genetic events (15–17). In particular, Armitage and Doll (17) examined human age-specific mortality rates, $OBS(h, t)$ for several cancer sites and noted a relationship, limited to the age range of 25–75, of the form $\log OBS(h, t) = \log K + (n-1) \log t$, where “*n*” was the number of required oncogenic events in a constant number of cells at risk.

This treatment, presented as a hypothesis by the authors, suggested that *n* might be 5, 6, or 7, the source of the idea that such numbers of oncomutations are required in human carcinogenesis (17).

Two-stage models

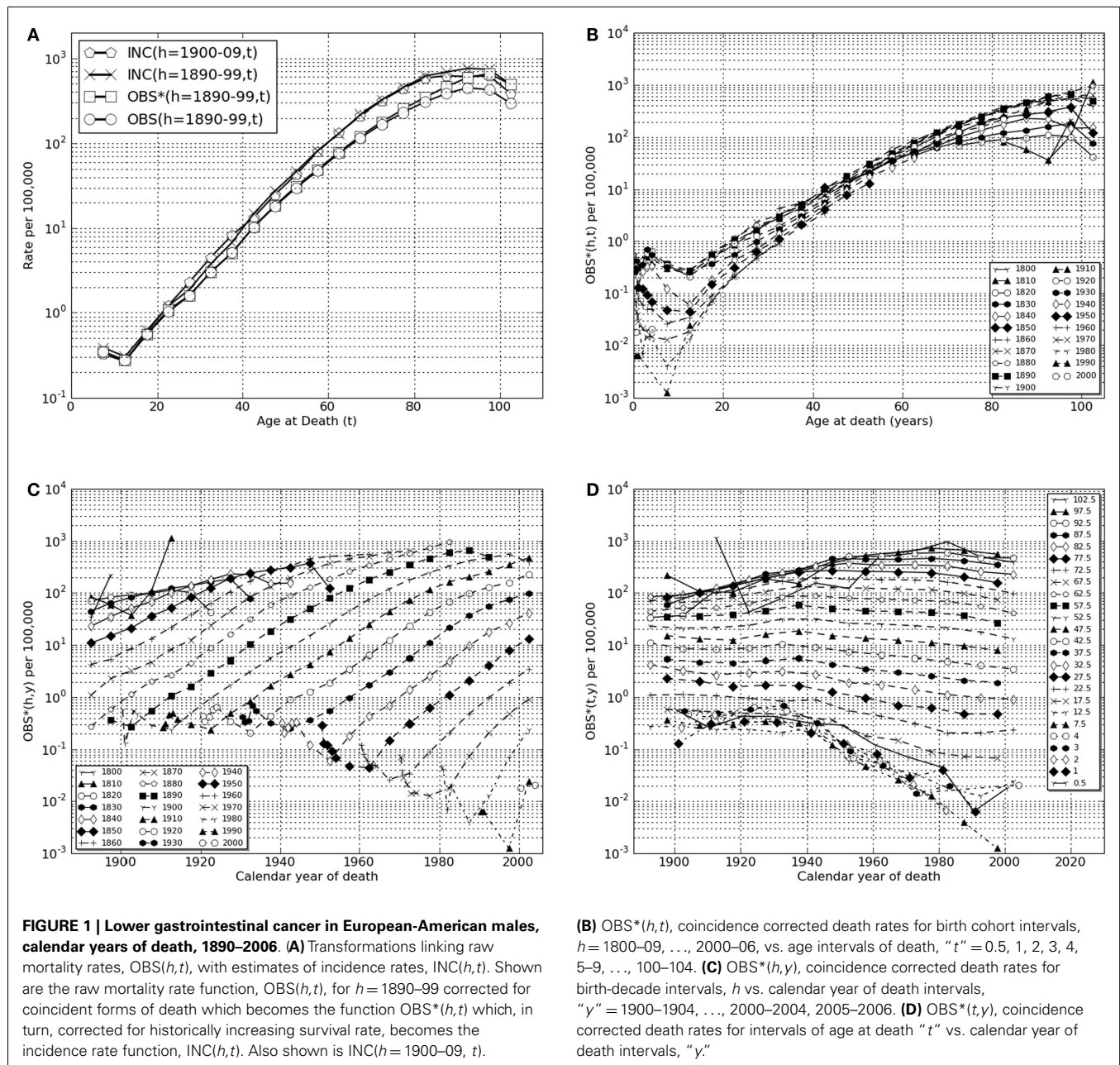
However, the “one-stage” models did not account either for growth of stem cell numbers in development or for clinically observed slowly growing preneoplastic lesions such as colonic adenomas. Platt was apparently the first to ask if early mutations could create a cell population with a growth advantage in which later necessary oncomutations could occur (18). Armitage and Doll responded to Platt’s suggestion by testing a model in which a single normal “cell at risk” was *initiated* by “*n*” required events could give rise to an exponentially growing preneoplastic colony in which any “cell at risk” was *promoted* to a tumor forming cell by “*m*” independent events (1). They chose the words *initiation* and *promotion* from the findings in experimental animal studies that certain chemical treatments could create a latent capability of tumor formation (initiation) that subsequent chemical treatments could provoke into growth of a visible neoplasia (promotion) (19). Their reasoning may be represented as $\log OBS(t) = \log Le^{-\mu t} = \log L + \mu t$ where *L* is a constant related to the product of initiation and promotion mutation rates and μ is the annual growth rate of preneoplastic cells at risk in adult tissue. Initiation events were posited to occur at constant annual rates per person from birth through old age. Human populations were considered to be homogeneous with regard to risk. It was noted that initiated cells at risk “had sufficient selective advantage to double in number about every 5 years” (1). They noted that both the “one-stage” and “two-stage” models could be fit by judicious choice of parameter values to age-specific mortality data for the age of death interval 25–75 years and, therefore, they could not exclude either possibility on this basis alone.

Limitation of “two-stage” model: declining cancer rates in extreme old age

However, neither the one-stage nor two-stage cancer models accounted for age-specific rates outside of the interval of 25–75 years. The rates of most forms of cancer decline from the first years of life to a minimum in mid-juvenile years, increase sub-exponentially into old age, reach a maximum then decline in extreme old age (Figure 1) (14, 20). Doll’s group ascribed the apparent maximum raw mortality rate in old age to either under-diagnosis of cancer in old age or the presence of a subpopulation that was not at risk (21). We algebraically amended the basic two-stage model using the suggestion of to account for a subfraction of the population, *F*, could be at risk for all of the events required for carcinogenesis in a particular organ, while the subfraction $(1 - F)$ was not at risk. This modification resulted in improved fits of prediction to observed incidence rates for colorectal cancer in the United States in old and extreme old age (14, 20). Predictions even more closely related to cancer incidence data may be expected if explicit partitioning of the general populations into subpopulations differing in quantitative risk factors are introduced, e.g., oncomutation rates (6), preneoplastic growth rates, and even individual size at maturity. Other biological hypotheses exist. For instance, it has been suggested that senescence itself suppresses preneoplastic or neoplastic growth (22, 23).

Initiation

Kinds and number of initiation mutations, *n*. More than 50 years after Armitage and Doll’s pioneering postulates, the



number, nature, and origins of oncomutations that *initiate* the slow growth of preneoplastic lesions and later *promote* a preneoplastic cell to the rapid growth of neoplastic lesions remain obscure for most common cancers, e.g., lung, breast, prostate, pancreas. For a few common cancer types, initiation events have been convincingly associated with the loss of both maternal and paternal alleles of specific autosomal tumor suppressor genes, e.g., colorectal (*APC*), kidney (*VHL*), nervous system (*NF1*, *NF2*), and basal skin (*PTCH*) cancers (24). Accordingly, we chose colorectal or “lower GI tract” cancer as a first example. The inherited early onset syndrome of multiple adenomatous polyps in the lower GI tract (colon and rectum), familial adenomatous polyposis coli or FAP, usually arises from inherited heterozygosity of the *APC* gene.

Loss of the second *APC* allele represents a necessary and possibly sufficient event for *initiation* followed by appearance of a growing adenomatous polyp as a preneoplastic lesion, i.e., the number of required colonic initiating mutations is equal to, or greater than, 2.

The deductions of Knudson (25) based on observation and analysis of certain inherited pediatric cancer syndromes such as Wilm’s tumor and retinoblastoma had previously indicated that two mutations in an organogenic lineage were necessary and sufficient to cause these diseases. A two-stage cancer model explicitly positing $n = 2$ was subsequently developed (26). Others motivated or dismayed by the large number and kinds of genetic and epigenetic changes displayed by most, but not all, cancers have

postulated that a much greater number of successive events are required (27, 28).

Limitation to the period of organ development. Influenced by studies and models of mutation in bacterial populations (10) two cancer modelers independently noted that the age-specific exponential increase in cancer rates in adults might have origins in the approximately exponential increase in organ cell number and, thus, initiated cell number from the beginning of organogenesis to maturity (29–31). A two-stage model of colon cancer incorporating mutation in juvenile (but not fetal) growth periods yielded expectations in concordance with observation (14, 20). A general, untested assumption that fetal mutation rates would/should be very low seems to have blocked further examination of the possibility that fetal mutation was a strong driver of tumor initiation. This idea was dispelled by direct assay at the DNA level of point mutations in micro-anatomical samples of upper bronchial epithelium from 15 healthy human lungs (6, 32).

In this present effort addressing mutations and cancers of the colon we have had recourse to literature reports from which it is possible to obtain reasonable first estimates of fetal/juvenile mutation rates in development of that organ. For instance, patients heterozygous for the *APC* gene are reported to have thousands of adenomatous polyps in adulthood presumably resulting from a single mutation inactivating the single active inherited *APC* allele. The inherited condition appears to be fully penetrant: all known *APC* heterozygotes have displayed multiple adenomas and, if untreated, adenocarcinomas. It should also be noted that the number of adenomatous polyps varies widely within family members afflicted with FAP, an observation extended to mutation rates in human lung epithelium that were found to vary about sevenfold from upper to lower decile estimates (6, 11, 24, 32).

Here we regard polyp number as a probable underestimate of actual mutated stem cells as many patients die before polyps arising late in juvenile development could be recognized. Allowing for 5000–10,000 potential polyps from the stem cells forming about 10^7 colonic crypts suggests that events per stem cell doubling (~ 23.25 to create the crypt number observed) would be of the order of $2\text{--}4 \times 10^{-5}$ gene-inactivating events per stem cell doubling for the *APC* gene. These events would include point mutations, chromosomal, and other possibilities including reciprocal recombinations. Our derived estimate of the geometrical mean rate of loss of the second allele of *OAT* in colons of heterozygotes is also $\sim 2\text{--}4 \times 10^{-5}$ (12). We later use these estimates to compare initiation mutation rates predicted for different possible numbers, “*n*” of required initiation mutations.

Nature of cells in which initiation mutations occur. Confusion arises in the use of these models in that the earliest efforts treated the number of “cells at risk” of initiation or promotion as constant (15–17) while the 1957 effort of Armitage and Doll treated initiation as occurring in a constant number of cells at risk while promotional mutations occurred in a growing population of initiated cells (1). Normal “cells at risk” might have included only stem cells or any of the stem plus non-stem cells of an epithelial maintenance turnover unit (14, 20). Here we explore the possibility that normal “cells at risk” are solely metakaryotic stem cells

(3, 4) that may be readily detected and counted in growing tissues, preneoplastic, and neoplastic lesions. Of course, there are other possibilities, the study of human organogenic and carcinogenic stem cells is still in its own gestation period.

Promotion

Clinical observations have shown that preneoplastic lesions of several organs continue to grow after maturity and give rise to lethal neoplasias through extreme old age. Promotional genetic changes or other events, if any, may therefore be hypothesized to occur throughout adult life in preneoplastic colonies in either symmetric (net growth) or asymmetric (turnover) amitoses of initiated metakaryotic stem cells. Uncertainties abound regarding the process called “promotion.”

First, it must be emphasized that no genetic or epigenetic events have yet been found that fulfill the logical requirements for a rare human promotion event. Secondly, as pointed out by Peto, mathematical analyses of non-linear age-specific incidence functions alone cannot define the number of “*n*” initiation or “*m*” promotion mutations using Armitage and Doll’s one-stage or two-stage models (33).

Progression

The steps toward deaths from primary neoplasias such as colonic adenocarcinomas and/or their metastases may involve one or more additional mutations/events in a stem cell during primary tumor growth. Here we accept the approximation of ~ 2.5 years between the last required promotion event and death (1).

USING MORTALITY AND SURVIVAL DATA TO ESTIMATE AGE-SPECIFIC INCIDENCE, $INC(h,t)$

U.S. age-specific cancer mortality rates (1900–2006)

Trial of any model of age-specific cancer experience requires a robust set of data to which predictions may be compared. The MIT-administered website¹ provides researchers with age-, gender, ethnicity-, and birth cohort-specific population and mortality data organized as Excel® files with several different forms of tabular and graphic summaries. Data are provided as recorded annually from 1900 through 2006 in the USA for most common cancer types and other major forms of mortality such as cardiovascular and cerebrovascular diseases. (Data for Japan are provided for most forms of cancer from 1952 through 1996.) These data permit examination of disease-specific mortality rates for the entire United States throughout the adult lifetimes (15–104 years) of persons of the same birth year or decade cohort born from ~ 1885 to 1910. Earlier birth decades lack data that were unrecorded before 1900. Later birth decade cohorts have not yet reached extreme old age.

Many forms of cancer such colorectal, pancreatic, and brain cancers as well as vascular diseases and Type II diabetes display a sub-exponential increase in mortality rates from maturity to a maximum then declining rate in extreme old age. The calculated model presented here, $CAL(h,t)$ is aimed at these forms of disease. (However, cancers of other organs including pharynx, tongue,

¹<http://mortalityanalysis.mit.edu>

testes, ovaries, uterus, and breast display a clearly different age-specific mortality rate function in which a distinctly more rapid rise in age-specific mortality rates are observed in young than in middle to older aged adults.) Data for many cancer sites are further limited because U.S. national recording began later than 1900, e.g., lung or pancreas in 1930.

This database is updated using ASCII files from the National Center for Health Statistics. U.S. data is organized by gender and major ethnic groups: European Americans and Non-European Americans, the latter comprising primarily African Americans. Information for years 1900–1991 were manually transcribed and re-organized from successive volumes of Vital Statistics of the United States published by the U.S. Census Bureau (1900–1935) and then by the U.S. Public Health Service (1936–1992) (Pablo Herrero-Jimenez), from digital versions of the same publication (1992–1997) (Efren Gutierrez) and directly from comprehensive ASCII files since 1998 (Karl Rexer, Rebecca Kusko, Lohith G. Kini). The computer program Cancer Fit v.5.0 described below interfaces with downloaded Excel® files from this data set.

As originally organized by the U.S. Census Bureau, the number of deaths for each form of mortality were recorded in each calendar year, y , for cohorts defined by gender, ethnicity, and age of death intervals, $t = 0, 1, 2, 3, 4, 5-9, \dots, 100-104$ years. In 1900 the two major ethnic designations were “white” and “non-white.” As “white” included families immigrating from Europe or through former Spanish, French, and Portuguese colonies of North and South America the classification “white” was denominated as European-American [EAM (male), EAF (female)] and the “non-white” consisting primarily of families of African descent was designated Non-European-American [NEAM (male), NEAF (female)].

Transforming raw mortality data to incidence estimates

Any single recorded death included the (a) recorded primary cause of death, (b) gender, (c) ethnicity, (d) age at death, “ t ” (e) year of death, “ y ,” and allowed calculation of (f) year of birth, “ h .” Age-specific mortality from a defined disease effecting a specific gender and ethnic group here characterized by the three related numbers, year of birth (h), age at death (t), and calendar year of death (y) such that $t = y - h$. Biases such as misdiagnosis are recognized but beyond the scope of this treatment.

Thus the data of annual number of deaths for lower GI tract cancer mortality in European-American males from 1900 to 2006 may be represented as MOR (EAM, lower GI tract, y, t, h). Since this paper will use only data for lower GI tract cancer in the EAM group we will shorten this designation to the form MOR(y, t, h).

It is possible to group data entries into defined intervals of calendar birth or death years or age of death as desired. Herein group data from age of birth decades such that the grouped variable, h , has values of 1800–1809, 1810–1819, \dots , 1990–1999, 2000–2006 (see text footnote 1). Similarly the calendar years of death may be grouped such that the variable representing grouped calendar years of death, y , has values of 1900–04, 1905–09, \dots , 1995–1999, 2000–2004, 2005–06. And the grouped variable for age of death, t , has values of 1, 2, 3, 4, 5–9, 10–14, \dots , 95–99, 100–104.

For studies of age-specific cancer rates, the data are organized as birth year cohorts and age of death intervals as, for instance,

MOR(h, t) in which the grouped mortality data for each birth-decade cohort, e.g., 1890–99 would be plotted as a function of age of death intervals in adult life, $t = 15-19, \dots, 100-104$.

The raw data from which all other variables are derived are the number of deaths reported each year, MOR(h, t), and the reported population size, POP(h, t). Their ratios, OBS(h, t) serve as a raw estimate of age-specific (t) death rates (deaths per person alive within the age interval) for each defined birth cohort (h).

However, historically high death rates from other causes among the youngest and oldest members of the population automatically lead to significant underestimation of persons who would have died of the disease observed because they died within the interval of another cause. To overcome this bias we needed a correction factor for each age of death interval of each birth cohort. To provide this we defined the set of values for deaths by all causes (total deaths) in each age of death interval as TOT(h, t) = MOR[all causes, (h, t)/POP(h, t)]. We reasoned that a first order correction for coincident forms of death could be accomplished by defining a coincidence corrected age-specific mortality rate, OBS*(h, t):

$$\text{OBS}(h, t) = \text{MOR}(h, t) / \text{POP}(h, t) \quad (1)$$

$$\text{TOT}(h, t) = \text{MOR}(\text{all causes})(h, t) / \text{POP}(h, t) \quad (2)$$

$$\text{OBS}^*(h, t) = \text{OBS}(h, t) / [1 - \text{TOT}(h, t) + \text{OBS}(h, t)] \quad (3)$$

In **Table 1** we show the exact steps used to define OBS*(h, t) for the birth-decade cohort $h = 1890-99$ for the age of death interval $t = 100-104$ years for lower GI tract cancer in European-American males. Here it may be seen that the estimate of OBS*(h, t) = 501×10^{-5} is derived as shown from 10 separate estimates of the number of recorded deaths, MOR($h = 1890, \dots, 1899, t = 102$) and corresponding 10 separate population values, POP($h = 1890, \dots, 1899, t = 102$) corrected for coincident deaths within each year by the values of TOT($h = 1890, \dots, 1899, t = 102$).

Accounting for historical improvements in colorectal cancer treatment

Given the values of OBS*(h, t) one must account for the fact that a correctly diagnosed tumor of the lower GI tract would not necessarily be mortal. There is convincing evidence that the probability of surviving at least 5 years after competent diagnosis, SUR(h, t), has been rising over the last ~70 years (14, 20). The values of OBS*(h, t) derived from deaths alone must represent an underestimate of the age-specific disease incidence, defined here as the fraction of persons at a given age in which the disease first appears, INC(h, t). If the values of SUR(h, t) are available, the estimate of disease incidence may be improved by calculating INC(h, t) = OBS*(h, t)/[1 – SUR(h, t)]. This step is illustrated in **Table 2** in which the values of OBS*(h, t) are transformed into values of INC(h, t) by use of historical 5-year post-diagnosis survival estimates for lower GI tract cancer (14).

$$\text{INC}(h, t) = \text{OBS}^*(h, t) / [1 - \text{SUR}(h, t)] \quad (4)$$

Estimation of survival rates especially in old age continues to be a source of considerable uncertainty. Here we used the values

Table 1 | Arithmetic steps in definition of $OBS^*(y = 1992, \dots, 2001; t = 100-04 \text{ years})$, from reported deaths, $MOR(y, t)$ for lower GI tract cancer and population sizes, $POP(y, t)$ for EAM.

Year of death (y)	Year of birth (h)	$MOR(h = 1890-99, t = 100-104)$	$POP(h, t)$	$TOT(h, t)$	$OBS(h, t) = MOR(h, t) / POP(h, t) \times 10^5$	$OBS^*(h, t) = OBS(h, t) / (1 - TOT(h, t) + OBS(h, y) \times 10^5$
1992	1890	27	4656	0.44	580	1035
1993	1891	19	6281	0.35	303	466
1994	1892	19	6484	0.36	293	458
1995	1893	18	6847	0.36	263	411
1996	1894	24	7273	0.34	330	500
1997	1895	20	7642	0.36	262	409
1998	1896	25	7682	0.32	325	478
1999	1897	22	7719	0.32	285	419
2000	1898	23	7753	0.32	238	350
2001	1899	30	8272	0.29	363	511
1992–2001	1890–1899					501

This is the age-specific death rate corrected for coincident deaths in the same reporting years, y . Applies eqs 1, 2, and 3. Note that these data comprise the set used to define the value of $OBS^*(h = 1890-99, t = 100-04)$ for the birth-decade cohort of 1890–99 dying from colorectal cancers in the 100–104 age interval, the mid year age being 102 years: $1992-102 \Rightarrow 1890$, $1993-102 \Rightarrow 1891$, ..., $2001-102 \Rightarrow 1899$.

Table 2 | Arithmetic steps in definition of $INC(h, t)$, the age-specific incidence rate for cohort h corrected for historically improving 5-year survival rates, $SUR(h, t)$, shown here for lower GI tract cancer for two birth-decade cohort intervals of birth, $h = 1890-99$, and $h = 1900-09$. Applies eq. 4.

Age at death interval, t	$OBS(h, t)$	$OBS^*(h, t)$	$SUR(h, t)$	$INC(h, t)$	$OBS^*(h, t)$	$SUR(h, t)$	$INC(h, t)$
	$h = 1890-9$	$h = 1890-9$	$h = 1890-9$	$h = 1890-9$	$h = 1900-9$	$h = 1900-9$	$h = 1900-9$
15–19	5.49E–06	5.49E–06	0.1	6.10E–06	4.96E–06	0.15	5.84E–06
20–24	1.02E–05	1.07E–05	0.1	1.19E–05	9.64E–06	0.2	1.21E–05
25–29	1.58E–05	1.59E–05	0.15	1.87E–05	1.72E–05	0.25	2.29E–05
30–34	3.02E–05	3.00E–05	0.2	3.76E–05	3.15E–05	0.3	4.50E–05
35–39	5.12E–05	5.09E–05	0.25	6.78E–05	5.54E–05	0.33	8.27E–05
40–44	1.02E–04	1.03E–04	0.3	1.47E–04	8.88E–05	0.33	1.33E–04
45–49	1.81E–04	1.83E–04	0.33	2.74E–04	1.49E–04	0.4	2.48E–04
50–54	2.99E–04	3.12E–04	0.33	4.65E–04	2.57E–04	0.4	4.28E–04
55–59	4.74E–04	4.92E–04	0.4	8.20E–04	4.46E–04	0.44	7.96E–04
60–64	7.61E–04	7.87E–04	0.4	1.31E–03	7.48E–04	0.44	1.34E–03
65–69	1.16E–03	1.22E–03	0.44	2.18E–03	1.18E–03	0.44	2.11E–03
70–74	1.69E–03	1.80E–03	0.44	3.21E–03	1.76E–03	0.44	3.14E–03
75–79	2.30E–03	2.52E–03	0.44	4.49E–03	2.44E–03	0.44	4.36E–03
80–84	3.04E–03	3.44E–03	0.44	6.15E–03	3.27E–03	0.44	5.84E–03
85–89	3.82E–03	4.60E–03	0.33	6.86E–03	4.14E–03	0.44	7.39E–03
90–94	4.46E–03	5.91E–03	0.22	7.58E–03	4.72E–03	0.33	7.04E–03
95–99	4.24E–03	6.56E–03	0.11	7.37E–03	5.51E–03	0.22	7.06E–03
100–104	2.93E–03	4.91E–03	0	4.91E–03	3.85E–03	0.11	4.33E–03

recorded in various clinical studies for American males up to age 80–84 (14) and estimated subsequent values as a linear decline to $SUR(h, t) = 0$ at age 100–104 based on discussions with clinicians about their experiences with declining use of surgery and/or other attempts at therapy in the extremely aged. The estimates of $SUR(h, t)$ used for the birth cohorts of 1890–99 and 1900–99 are listed in Table 2 so that readers may understand this process and that our estimates of $SUR(h, t)$ for ages 85–104 are “best guesses.”

Test cohort: EAM lower GI tract cancer mortality in the U.S., 1900–2006

Deaths from cancer of the lower gastrointestinal tract (present ICD9 codes: 152, 153, 154) are predominantly from colorectal cancer (see text footnote 1). Figure 1A uses a semi-log plot to illustrate the transformations from raw mortality rates to incidence rates: from $OBS(h = 1890-99, t)$ to $OBS^*(h = 1890-99, t)$, then to $INC(h = 1890-99, t)$. Also shown is the result of the same transform of the data of the birth cohort of the succeeding decade,

INC($h=1900-09, t$), which yielded similar estimates of for all values of “ t ” (Table 2).

In Figure 1B the coincidence adjusted mortality rate, $OBS^*(h, t)$ is presented as a function of age of death, t , on a semi-log scale for each of successive birth cohort intervals, “ h ,” from 1800–09 through 2000–06. These data demonstrate that colorectal cancer death rates in older adults ($t > 65$) rose throughout the birth cohorts of the nineteenth century reaching an approximately stable maximum in and after the birth cohort of 1880–89 modified to the extent discussed above by improvements in medical treatment represented by increasing values of $SUR(h, y)$.

In Figure 1C these same data for colorectal cancer are plotted so that changes $OBS^*(h, y)$ as a function of age of death interval, “ t ” may be seen as functions of historical year of death interval “ y .” Here one may clearly see the historical changes in successive birth cohorts as maximum mortality rates rose for unknown reasons in cohorts born in the early nineteenth century and declined steadily for cohorts living in the twentieth century when improved surgical procedures increased survival rates.

In Figure 1D the data are plotted as $OBS^*(y, t)$ vs. “ y ” to illustrate the historical changes within each age of death interval, “ t ” as a function of historical year of death interval, “ y .” This form of presentation shows a significant decrease in older adult death rates ascribable in whole or part to advances in medical practice after 1950 that increasingly effect each successive birth cohort (Table 2) and a marked, previously unrecognized and unexplained decrease in pediatric death rates dating from ~1940.

AMENDMENTS OF THE ARMITAGE–DOLL “TWO-STAGE” CANCER MODEL

Limitation of initiation mutations to the fetal/juvenile stem cell doublings

Growth of normal fetal/juvenile stem cells is here modeled as a series of “ a ” net binomial doublings ($a = 0, 1, 2, \dots, a_{\max}$) defining the growth of the number of stem cells in an organ, $N(a)$, throughout fetal and juvenile growth to maturity. This does not mean that we assume that each and every fetal/juvenile cell survives and grows exponentially by binomial fissions. We are aware that many organs, e.g., lung, breast, prostate, vascular system, grow as arborated ductal structures. But we note that their net growth may be reasonably approximated as a binomial expansion. Our mutation rates are in essence the probability that any new cell in the stem cell expansion resulting in a mature organ has newly acquired a required oncomutation (event) in the net cell doubling interval represented by “ a .”

The number of fetal stem cells during growth $N(a)$ is thus simply represented as $N(a) = 2^a$. Initiation is postulated to occur in any stem cell by acquisition “ n ” required initiation mutations, i, j, \dots, n , occurring in any order at constant mutation rates R_i, R_j, \dots, R_n per doubling (14, 20). The expected number of newly initiated stem cells in each doubling period “ a ,” $NEW_{\text{init}}(a)$ may be expressed as:

$$NEW_{\text{init}}(a) = n \left(\prod_i R_i \right) a^{(n-1)} 2^a \quad 0 \leq a \leq a_{\max} \quad (5)$$

In the fetal/juvenile model organogenic stem cells are posited to reach maturity represented by “ a_{\max} ” doublings with high constant

mutation rates and to undergo metamorphosis to maintenance stem cells with no net additional net cell growth and much lower mutation rates (6).

Assuming each of the $\sim 10^7$ adult colonic crypts to be represented at juvenile/adult metamorphosis by a single metakaryotic stem cell, the number of net doublings at maturity, a_{\max} , is about 23.25, i.e., $10^7 \sim 2^{23.25}$ (14, 20). The metakaryotic mutator/hypermutable stem cell lineage of human organ anlagen is here formally postulated to begin in gestational week 4–5 with creation of two metakaryotic stem cells from symmetrical amitosis of a single precursor embryonic mitotic stem cell at $a = 0$ (4). At birth, we estimate that a human colon contains $\sim 2^{20}$ colonic crypts each containing a basal metakaryotic stem cell; thus at birth, $a \sim 20$, at maturity, $a = a_{\max} \sim 23.25$.

Promotion mutations during preneoplastic stem cell doublings

While the age of the developing organ may be designated as “ a ,” each initiated stem cell arises as a single cell and is here posited to double in parallel with its organogenic lineage till a_{\max} and then continue to grow in adult tissue at a rate of doubling, μ , similar to that of the juvenile organ in which it resided (14). At maturity, initiated colonies arising early in fetal growth will be larger than those initiated at the last pre-maturation doubling that would consist of a single initiated stem cell. (30, 31) Each initiated colony is posited to grow at a constant rate, μ , until one preneoplastic stem cell experiences “ m ” required events to promote it to a neoplastic stem cell that founds a rapidly growing, potentially lethal adenocarcinoma with sequelae, including metastasis, leading to death. Here we adopt the suggestion of Armitage and Doll (1) that the time between promotion and death is about 2.5 years in adults.

Transforming stem cell doublings, g , into human age, t

In order to express the idea that initiated colonies arise throughout fetal/juvenile growth of organogenic stem cells and continue to grow in adult tissues we needed to define a continuous variable that expressed the age and, therefore the size, of each preneoplastic lesion throughout both fetal/juvenile and adult life.

To this end we introduce the continuous variable “ g ” that is equal to “ a ” for all values of “ a ” from $a = 0$ to $a = a_{\max}$ and then increases by one (1) for each adult preneoplastic stem cell doubling period in years represented by $1/\mu$ where μ is the doubling rate of stem cells in an adult preneoplasia. Assuming an age of organ maturity as ~ 16.5 for males and a 2.5-year interval between a promotion event in an initiated preneoplastic stem cell and death, the relationship between age in years, t , and age in terms of stem cell doublings since the beginning of organogenesis when $a = g = 0$ is simply:

$$g = \mu (t - 16.5 - 2.5) + a_{\max} = \mu (t - 19) + a_{\max} \quad (6)$$

This introduction of the parameter “ g ” is essential to our modeling approach in that it provides a means to relate events driven by stem cell doubling intervals, “ g ,” to mortality rates recorded in age of death intervals in adult life, “ t .”

Assembling the elements into a single continuous model of cancer incidence

After initiation in any fetal/juvenile doubling “ a ” growth of pre-neoplastic stem cells as a colony is modeled as a series of “ $(g - a)$ ” binomial doublings [$(g - a) = 0, 1, 2, \dots$]. In each preneoplastic stem cell doubling events such as mutations may occur until any preneoplastic stem cell experiences “ m ” required promotion events (A, B, \dots, m). These events are posited to occur at constant mutation rates per doubling: R_A, R_B, \dots, R_m . The expected number of newly initiated stem cells in preneoplastic doubling period “ $(g - a)$ ” $NEW_{prom}(g - a)$, is simply:

$$NEW_{prom}(g - a) = m \left(\prod_m R_A \right) (g - a)^{(m-1)} 2^{(g-a)} \quad 0 \leq (g - a) \quad (7)$$

Under these assumptions the number of organogenic doublings “ a ” at initiation and the number of preneoplastic doublings “ $(g - a)$ ” after initiation sum to “ g .” In each organogenic-doubling interval “ a ” new preneoplastic colonies are created (initiated) and these colonies grow until promotion and subsequent death remove them.

For each organogenic stem cell doubling period “ a ” we now require an expression for the expected number of colonies per individual that are promoted in any interval of $(g - a)$ thereafter. Since the promotion of a preneoplastic stem leads to death the number of surviving preneoplastic after initiation in interval “ a ” will decline in each interval $(g - a)$.

For the set of initiated colonies arising in organogenic interval “ a ” the expected number of promotions in the subsequent lifetime intervals, g , $EXP(a \rightarrow g)$ may be expressed as:

$$EXP(a \rightarrow g) = n \left(\prod_n R_i \right) a^{(n-1)} 2^a d \left[1 - \exp \left(-m \left(\prod_m R_A \right) (g - a)^{(m-1)} 2^{(g-a)} \right) \right] / d(g - a) \quad (8)$$

While this expression may be computed for all values of “ n ” and “ m ” and is used in our computations for varying values of “ n ” and “ m ” below, we here introduce the restriction of $n = 2$ and $m = 1$ for the purpose of clearer illustration.

$$EXP(a \rightarrow g) = 2R_i R_j a 2^a d \left[1 - \exp \left(-R_A 2^{(g-a)} \right) \right] / d(g - a) = 2 \ln 2 R_i R_j R_A 2^g a \exp \left(-R_A 2^{(g-a)} \right) \quad (9)$$

Of course, the expected number of promotions to neoplasia in an adult in any interval “ g ” $V(a \rightarrow g)$, is simply the sum of expected promotions to neoplasia from the initiations of organogenic stem cells in each of the organogenic stem cell doubling intervals of $a = 0$ to $a = a_{max}$:

$$V(a \rightarrow g) = \sum (a = 0 \text{ to } a_{max}) 2 \ln 2 R_i R_j R_A 2^g a \exp \left(-R_A 2^{(g-a)} \right) \quad (10)$$

This process is illustrated in **Figure 2** in which the contribution to promotion at age “ g ” from initiation at each organogenic doubling “ a ” is shown to rise and fall with “ $(g - a)$.”

Figure 2 embodies our central argument that initiation is restricted to the metakaryotic stem cell doublings of organogenesis in the fetal/juvenile period.

The sum of these terms from initiations in all organogenic-doubling intervals “ a ” approximates well the observed lifetime incidence rate of many cancer types including colorectal cancer: it increases sub-exponentially, reaches a maximum in old age and declines appreciably in extreme old age. The earliest initiations of fetal organogenesis drive the tumor incidence rate of juveniles and young adults, the initiations of adolescent organogenesis drive the tumor incidence rate in extreme old age (31).

Under these conditions the expected number of newly promoted lesions through the end of any adult lifetime doubling interval “ g ,” $CAL(g)$, is:

$$CAL(g) = (1 - e^{-V(a \rightarrow g)}) \quad (11)$$

Additional possible modeling elements

Stratification of risks in the population. Stratification, or differences, within the population may encompass any environmental or heritable condition required for cancer death. Some environmental risks appear to operate through growth stimulation of preneoplastic colonies as opposed to induction of initiating or promoting oncomutations, e.g., persons exposed to sunlight as adults or adult cigarette smokers. Populations may thus be “stratified” on the basis of exposure to such preneoplastic growth-promoting agents.

Other risks may well be stratified including fetal/juvenile initiation mutation rates (6, 32), promotion mutation rates, growth of promoted neoplastic stem cells (tumor progression), and tumor invasion/metastasis. It is well to note that mortality data modeling provides information about the subpopulation at risk for a form of cancer but nothing about persons who are not at risk.

We have previously represented the fraction of the population in whom all of the potential conditions necessary for cancer death are present as “ F ” imagining that a person either is or is not at risk for all necessary oncogenic processes. The corresponding fraction in which one or more necessary condition(s) of risk is absent has been represented as $(1 - F)$ (14, 20). Stratification need not, however, be an “all or none” phenomenon. Children grow to different sizes, which may create stratification with regard to the maximum number of stem cells at risk of initiation $2^{a_{max}}$. Children also grow at different rates and so may the preneoplastic lesions in different persons. We have reported stratification with regard to mutation rates in fetal/juvenile growth for both mitochondrial and nuclear genes (9). In progress is an effort to incorporate stratification with regard to initiation and promotion oncomutation rates and the growth rates of preneoplasias. Here we continue to employ “ F ” as an approximation for more precise stratification variables. Equation 11 rewritten to account for stratification in this way creates the model:

$$CAL(g) = F(1 - e^{-V(a \rightarrow g)}) / [F + (1 - F)e^{\int V(a \rightarrow g) dg}] \quad (12)$$

Competing forms of risk potentially affecting age-specific cancer rates. Epidemiological observations have also demonstrated that forms of cancer may share environmental or inherited risk factors

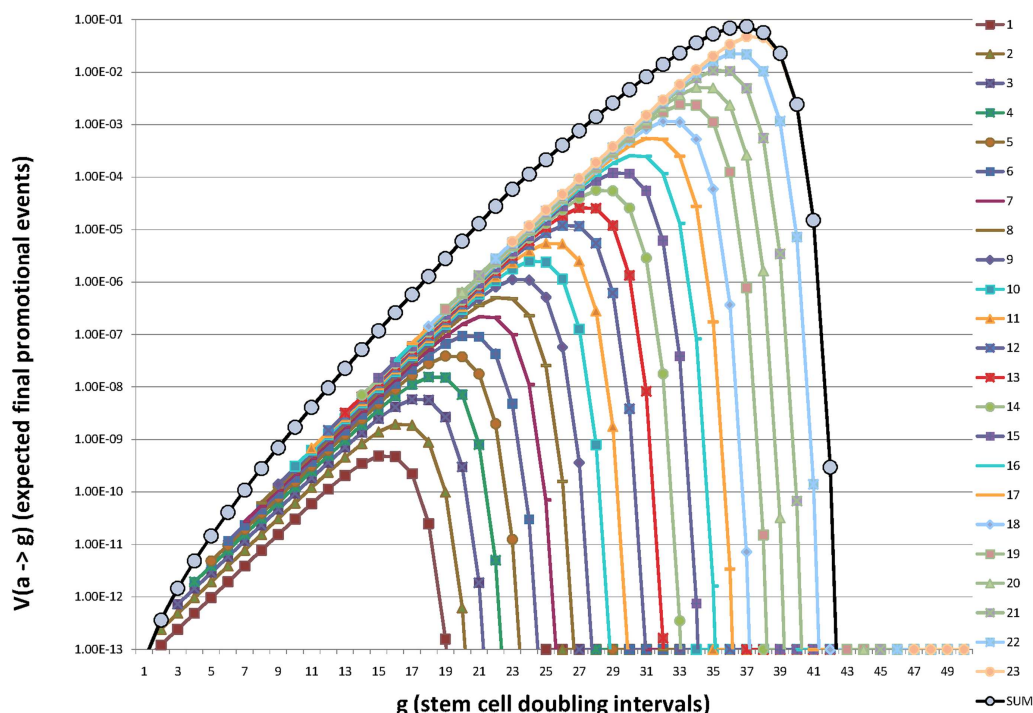


FIGURE 2 | Illustration of how initiation of organogenic stem cells in each successive stem cell doubling during fetal/juvenile growth to maturation contributes additively to the expectation of promotion (and death) throughout life. Here the beginning of organogenesis is represented as $a = g = 0$, birth is at approximately, $a = g = 20$, and maturity is at $a = a_{\max} = g = 23.25$. For a preneoplastic stem cell growth rate of $\mu = 0.2$ the maximum value of $V(a \rightarrow g)$ is reached at $g = 37$ or about or about

$t = 90$ years and $g = 40$ corresponds to $t \sim 105$ years. Note that earliest initiations presage fetal, juvenile, and young adult deaths. Initiations late in organ development account for deaths in extreme old age. Note the expected number of promotions (deaths), $V(a \rightarrow g)$ increases sub-exponentially from $g = 23$ to a maximum at $g \sim 37$ and then declines rapidly. In this example $g = 40$ represents ~ 105 years of age. Values used in this illustration were $n = 2$, $m = 1$, $R_i = R_j = 2.2 \times 10^{-5}$, and $R_A = 4.4 \times 10^{-5}$.

with another, e.g., breast and ovarian cancers, in which the death rates increase synchronously with age (see text footnote 1). The term “ f ” has been previously introduced to represent the fraction of persons that die of the observed cause among the set of mortal diseases with shared risks and synchronous changes in death rates (14, 20). We considered it possible that vascular disease death rates that rise synchronously with cancer death rates might represent a major form of competing risk. However, inspection of the stable and then declining death rates from cerebrovascular and cardiovascular death rates 1900–2006 while colorectal cancer rates have increased to stable maxima (accounting for increased survival) indicates that risks for vascular disease and colorectal cancer are probably not shared (see text footnote 1). Herein, treating colorectal cancer specifically, we have assumed *pro tempore* that there are no synchronously competing forms of death with shared risks for colorectal cancer (34), i.e., $f = 1.0$. As noted, this assumption is not valid for other forms of cancer such as breast, uterine, and ovarian cancers, which appear to share risk factors.

A COMPUTER PROGRAM TO ESTIMATE PARAMETERS OF CARCINOGENESIS IN THE TWO-STAGE MODEL ADJUSTED FOR FETAL/JUVENILE INITIATION: CANCERFIT V5.0

Using CancerFit v5.0

In order to permit cancer clinicians and biologists to explore quantitative hypotheses linking a wide range of biological and

population parameters to observed lifetime incidence rates the computer program CancerFit v5.0 was developed and is available without cost from the corresponding author or by download from <http://mortalityanalysis.mit.edu>. Any computer carrying MatLab®v.7.14 can run this program. Data for any form of cancer downloaded from <http://mortalityanalysis.mit.edu> as $OBS^*(h, t)$ and be directly imported into the program for the desired country (U.S. or Japan), gender, ethnic group, and birth year interval. Values of the parameters $(\Pi R_i)^{1/n}$, $(\Pi R_A)^{1/m}$, μ , and F may be either fixed or permitted to range with a chosen number of iterations for computation within each parameter range chosen by the researcher. On a typical 2009 computer such as a MacIntosh G5® 10^9 iterations of parameter combinations can be run in about 4 h. In typical use the combinations of variable iterations may range from 20^4 to 20^5 (1.6×10^5 to 3.2×10^6) requiring a few minutes for computation. The results of a single computation run are presented as the set of 100 best fits of the stipulated model to the values of $INC(h, t)$ as defined below.

The researcher may then inspect the ranges of values of all five parameters that together result in calculated age-specific incidence estimate $CAL(h, t)$ in good agreement with observation, $INC(h, t)$ for the stipulated values of “ n ,” “ m ,” and “ a_{\max} .” The tabulated results also provide the estimated parametric values for the fraction of persons initiated, F_{init} , and the area computed under the function $CAL(h, t)$, $AREA(h)$, from g (or t) = 0 to ~ 150 years,

which roughly estimates the expected (potential) number of lethal events from the observed disease per person over the lifetime of the cohort. Graphic representations are provided comparing $CAL(h, t)$ to $INC(h, t)$ for any desired fit in terms of linear or log values of $INC(h, t)$ vs. linear or log values of age of death, “ t .” It is hoped that this program and the organized set of U.S. and Japanese mortality data for a large set of cancer sites (see text footnote 1) will enable cancer researchers and students to explore quantitative hypotheses about cancer etiology from both a historical and age-specific perspective for the genders and ethnic groups represented.

Ranges of parameter values employed in calculating values of $CAL(h, t)$ for colorectal cancer

The central premise of the fetal/juvenile mutator/hypermutable stem cell hypothesis is that initiation mutations occur before adulthood when the maximum number of juvenile stem cells, $2^{a_{\max}}$ is reached, i.e., $0 \leq a \leq a_{\max}$ (6, 35, 36). In CancerFit v5.0 the value of a_{\max} must be specified for each organ studied, here we used 23.25 as an estimate for the colon/rectum (14). The values of F must lie between 0.0 and 1.0; for computation of $CAL(h, t)$ 20 values, 0.05, 0.1, ..., 0.95, 1.0, were used for each. Combinations of $n = 1-5$ and $m = 1-5$ were independently explored. Values of μ , the pre-neoplastic annual net growth rate, ranged with 20 linear iterations from 0.1 to 0.3 bracketing estimates from the annual growth rates of juvenile body mass, ~ 0.16 (14, 20). Values of $(\Pi R_i)^{1/n}$ representing the geometrical average of required oncomutation rates in initiation and of $(\Pi R_A)^{1/m}$ representing the geometrical time-averaged required onco-event rates in promotion ranged from 10^{-9} to 10^0 with 100 geometric iterations. After these geometrical means were approximated, the values were bracketed more closely to asymptotically obtain the best fits for each combination of n and m tested.

Goodness of fit

The values of $INC(h, t)$ vary more than a thousand-fold among the 18 age of death intervals recorded from maturity to old age (Figure 1). Statistical comparison of $CAL(h, t)$ to $INC(h, t)$ required a term that gave equal weight to all 18 intervals. Such a term is $(\log_{10} CAL(h, t) - \log_{10} INC(h, t))$ (that is zero when the terms are identical, i.e., their ratio is equal to 1.0. The square root of the average of the square of these 18 terms is here employed as a goodness of fit parameter, $GOF(h, t)$.

$$GOF(h, t) = \left\{ \sum [\log_{10} CAL(h, t) - \log_{10} INC(h, t)]^2 / 18 \right\}^{1/2} \quad (13)$$

$GOF(h, t)$ is thus akin to a standard deviation of $(\log_{10} CAL(h, t) - \log_{10} INC(h, t))$ (averaged over all adult life. If $GOF(h, t) = 0.1$ then the 95% confidence limits of the average ratio of $CAL(h, t)/INC(h, t)$ would be $10^{-0.2}$ and $10^{0.2}$ or 0.63 and 1.58. If $GOF(h, t) = 0.03$ the 95% limits would be 0.93 and 1.07. A $GOF(h, t)$ of greater than 0.1 suggests a poor fit of the model with stipulated parameters “ n ” and “ m ” or marked errors in the values of $INC(h, t)$ arising from either sampling error or bias. Errors arising from small sample sizes do not contribute significantly to $GOF(h, t)$ when large U.S. European-American cohorts are studied. In the death intervals containing the fewest recorded deaths the

Table 3 | Matrix of goodness of fit calculated as $GOF(h, t)$ for $n = 1, 2, 3, 4, 5$ and $m = 1, 2, 3, 4, 5$ under the parsimonious conditions of homogeneous risk ($F = 1$) and no synchronous mortal diseases sharing risk factors with colorectal cancer ($f = 1$).

		n				
		1	2	3	4	5
m	1	0.066	0.085	0.098	0.128	0.094
	2	0.072	0.079	0.084	0.052	0.074
	3	0.064	0.089	0.078	0.055	0.089
	4	0.098	0.109	0.098	0.071	0.103
	5	0.088	0.107	0.114	0.087	0.110

Values of the geometric means of initiating mutation rates $(\Pi R_i)^{1/n}$ and promotion mutation rates $(\Pi R_A)^{1/m}$ were permitted to range from 10^{-9} to 10^0 and the range of μ was set at 0.1–0.3. The results illustrate the fact that values of n and/or m cannot be derived by simply fitting the non-linear function, $CAL(h, t)$ represented by the algebraic model to the data, $INC(h, t)$: multiple combinations of n and m result in values of $GOF(h, t)$ indicating reasonable correspondence of the model to the data.

number of lower GI tract European-American male cancer deaths recorded for the birth cohort of 1890–99 were 246 for $t = 15-19$ (1908–1917) and 1999 for $t = 100-104$ (1993–2002) respectively (see text footnote 1). Based on the assumption of a Poisson distribution of deaths per living person within each age of death interval, t , the $GOF(h, t)$ for two random samples of the same population would yield an expected $GOF(h, t)$ of about 0.03.

Given the various biases and sampling errors expected for measurements of $OBS(h, t)$, $TOT(h, t)$, and especially $SUR(h, t)$ defining the derived function $INC(h, t)$ this would be as good a fit as might be expected for any model. When the values of $INC(h = 1890-99, t)$ were compared to $INC(h = 1880-89, t)$ and $INC(h = 1900-09, t)$ over all 18 age of death intervals $GOF(t)$ was ~ 0.043 a number not much greater than the value ~ 0.03 expected by chance alone (Table 3).

RESULTS

APPLICATION OF THE MODEL TO AGE-SPECIFIC COLORECTAL CANCER INCIDENCE IN A SPECIFIC COHORT

First, the best fits of $CAL(h = 1890-99, 15 < t < 104)$ were calculated for the 25 combinations of $n = 1-5$ and $m = 1-5$ under the parsimonious conditions of homogeneous risk and no synchronous mortal diseases sharing risk factors with colorectal cancer ($F = 1, f = 1$). Values of $(\Pi R_i)^{1/n}$ and $(\Pi R_A)^{1/m}$ were permitted to range from 10^{-9} to 10^0 and the range of μ was set at 0.1–0.3. The complete matrix of results is provided in Table 3. For $n = 2$ (an appropriate biological value only if the loss of both copies of the APC gene were necessary and sufficient for initiation in most colorectal cancers) and $m = 1$ as default assumptions the $GOF(h, t)$ was 0.085.

Second, we considered the possibility that stratification of risks within the population such as varying oncomutation rates might account for the observed discordance ($F, 1, f = 1$). To do this we compared $CAL(h, t)$ to $INC(h, t)$ under the additional assumption of inhomogeneous risk(s), i.e., the parameter “ F ” representing a hypothetical fraction of the population at risk was allowed to

Table 4 | Complete matrix of $GOF(h,t)$ for $m = 1, 2, 3, 4, 5$ and $n = 1, 2, 3, 4, 5$ with the assumption of inhomogeneous risk, i.e., the parameter “ F ” representing a hypothetical fraction of the population at risk was allowed to range from 0 to 1 ($F < 1$) and no synchronous mortal diseases sharing risk factors with colorectal cancer ($f = 1$).

		n				
		1	2	3	4	5
m	1	0.038	0.043	0.039	0.041	0.046
	2	0.036	0.048	0.042	0.046	0.054
	3	0.043	0.042	0.046	0.042	0.043
	4	0.047	0.047	0.035	0.040	0.044
	5	0.051	0.050	0.064	0.047	0.045

Values of $(\prod R_i)^{1/n}$ and $(\prod R_i)^{1/m}$ were permitted to range from 10^{-9} to 10^0 and the range of μ was set at 0.1–0.3. Comparison with the values of $GOF(h,t)$ of **Table 3** illustrates that the assumption of population inhomogeneity results in better agreement between model and incidence data. Such inhomogeneity has been reported as an appreciable range of fetal juvenile mutation rates for the human lung bronchial epithelium (6).

range from 0 to 1. The value of 0.043 was the minimum $GOF(h,t)$ observed for $n = 2$, $m = 1$ and the concordance in the age range 80–104 was appreciably improved. The complete matrix of these results is provided as **Table 4**.

Thirdly, we considered the possibility of both population inhomogeneity and a competing synchronous mortal disease having genetic and/or environmental risks shared with colorectal cancer ($F < 1$, $f < 1$). This assumption did not, however, further reduce the values of $GOF(h,t)$ consistent with the assumption that other appreciable forms of mortality do not share risk factors with lower G.I. tract cancers (34).

Figure 3 graphically depicts the degree of concordance of the trial conditions: $F = 1$ (population homogeneity with regard to risk) and $F < 1$ (population inhomogeneity) with adult lifetime incidence data for lower G.I. tract cancer in European-American males born 1890–99 $INC(h,t)$. It should be noted that for the assumption $F = 1$ discordance between $INC(h,t)$ and $CAL(h,t)$ was greatest at $t > 75$ years where underestimation of colorectal cancer as a cause of death by as much as 30% has been suspected in extreme old age (14, 20, 21, 37, 38). Note that our tactic of converting the raw mortality rate estimate $OBS(h,t)$ to the coincidental death corrected estimate $OBS^*(h,t)$ would not account for under-diagnosis post mortem. However, when $CAL(h,t)$ accounts both for fetal/juvenile mutation limitation (6) and the possibility of risk stratification in the population ($F < 1$) fits are considerably improved (**Figure 3**).

USE OF OBSERVATIONS IN HUMANS TO ESTIMATE NUMBER OF EVENTS REQUIRED FOR INITIATION (n) AND PROMOTION (m) IN COLORECTAL CANCER

Comparison of the predictions of the amended Armitage–Doll two-stage model represented as $CAL(h,t)$ to the observed lifetime adult age-specific incidence of colorectal cancer in European-American males born 1890–99 or 1900–09, $INC(h,t)$ achieved a higher degree of concordance over the entire adult lifespan than has previously been reported for any previous model of human

cancer. This result is consistent with the hypothesis that tumor initiation mutations are limited to the stem cells of the fetal/juvenile period and that there is a significant degree of heterogeneity of risk factors in the population ($F < 1$). We have reported considerable variation in the population with regard to fetal/juvenile mutation rates in the human lung (6). However, as noted by Peto, no determination of the values of the number of initiating (n) or promoting (m) events is possible by this approach alone as evidenced by the good fits for all 25 combinations of n and m tried (**Table 4**) (33).

However, we could now apply several observations delimiting some of the key parameters and discover what, if anything, may be suggested here using the stipulations that $F < 1$, $f = 1$.

First, we applied our observation that juvenile colonic crypts each contain a single metakaryotic stem cell (4) and that the number of crypts in an adult male colon is about 10^7 or $2^{23.25}$ (14). This value we associated with the maximum number of juvenile colonic stem cells before maturation to adult colonic maintenance stem cells with a form of eukaryotic behavior, i.e., we estimated that $a_{\max} = 23.25$.

Second, we again noted that colonic fetal/juvenile mutation rate estimates of $\sim 2\text{--}5 \times 10^{-5}$ gene-inactivation mutations per juvenile stem cell doubling have been derived from observation of adenomatous polyps or mutations in the colon for the *APC* and *OAT* genes (11, 12). Independent of the value of m stipulated, the only value of n that resulted in such an initiation mutation rate range is 2. For $n = 2$ the estimated rate was $\sim 3.5 \times 10^{-5}$; for $n = 1$, $\sim 5.7 \times 10^{-8}$; for $n = 3$, $\sim 3.6 \times 10^{-4}$.

The provision of the fetal/juvenile stem cell mutation rate thus permitted the conclusion that $n = 2$ events. These could comprise loss of both maternal and paternal *APC* alleles in the numerically dominant colorectal cancer pathway (24).

Third, we previously noted the annual juvenile mass doubling rate of U.S. males is about 0.158, significantly lower than estimates of preneoplastic growth rates of 0.18–2.0 generated when it is assumed that $n = 2$ and $m = 1$ (14).

Fourth, we estimated loss of heterozygosity fractions for many different polymorphic markers in colon tumors are about 0.125 for a single alleles (14). With preneoplastic doubling rates estimated to be between 0.158 and 0.2 the number of years between male maturity and the age at which mortality rate reaches a maximum represents about 15 doublings. We here made a rough estimate from these values: $0.125/15 = 8 \times 10^{-3}$ gene deletion events per doubling. This represented a lower bound on gene-inactivation or loss rates in preneoplastic stem cells. If loss of either allele for an autosomal gene were a promotional event the rate would be twice this estimate or $> 1.6 \times 10^{-2}$.

As the values of m from 1 to 5 (holding $n = 2$) increased, the geometric mean of the promotion event (mutation) rates increased. For $m = 1$ the estimate was 2.2×10^{-5} similar to the estimated rate of promotion events; for $m = 2$, $\sim 1.9 \times 10^{-3}$; for $m = 3$, 7×10^{-3} ; for $m = 4$, 1.4×10^{-2} ; for $m = 5$, $\sim 2.1 \times 10^{-2}$.

As the values of m increased the estimated growth rate of preneoplastic stem cells decreased. A decrease in estimated preneoplastic doubling rates with increasing values of m was expected. For values of $m > 1$, the generating function for promotion would rise as $(g - a)^{(m-1)}$ as originally imagined by the earliest cancer

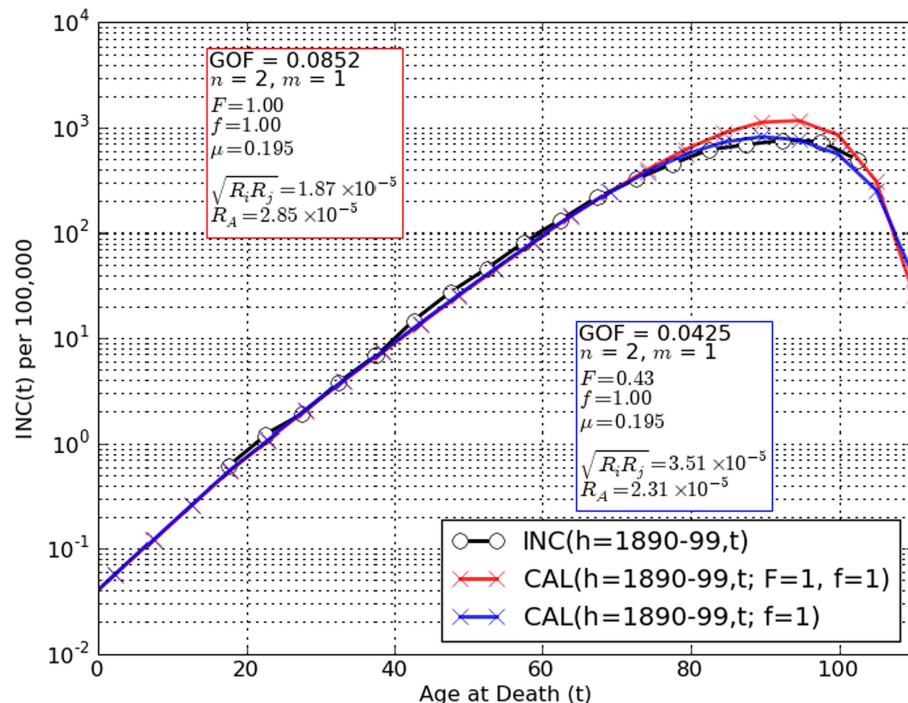


FIGURE 3 | Comparisons of observation $INC(h,t)$ (black circles) to $CAL(h,t)$ (red, blue crosses) for the two-stage cancer model amended for limiting tumor initiation mutations to the fetal/juvenile stem cells and $n=2, m=1$. A remarkably good fit is obtained for the assumptions that the population is homogeneous for risk factors such as oncomutation rates and/or preneoplastic growth rates ($F=1$) and there are no competing synchronous forms of death ($f=1$) [red crosses, $GOF(h,t)=0.085$].

However, a significantly lower value, $GOF(h,t)=0.043$ is observed when population risk stratification is assumed ($F<1$). Allowing for the possibility of an unknown competing synchronous form of death ($f<1$) did not decrease the values of $GOF(h,t)$ further. These findings are spurring further amendment of the two-stage model to use stratified variables for initiation and promotion mutation rates, preneoplastic growth rate and number of stem cells at maturity.

modelers (15–17). This supralinear term would force a decrease in the estimate of the exponential growth rate of preneoplastic stem cells, μ , in order to fit the data of $INC(h,t)$.

We noted that for $m=3$, $\mu=0.167$, and the geometrical mean of the three promotion event rates, $(\Pi R_A)^{1/m}$ was $\sim 7 \times 10^{-3}$. For $m=4$, $\mu=0.162$, and $(\Pi R_A)^{1/m}=1.4 \times 10^{-2}$. For $m=5$, $\mu=0.156$, and $(\Pi R_A)^{1/m}=2.1 \times 10^{-2}$. Given the several uncertainties in the estimate of preneoplastic gene losses and the range observed among losses, it seemed that values of $m=4$ or 5 would best fit the data. These observations further suggest that the growth rate of preneoplastic stem cells in males is indistinguishable from the juvenile growth rate of males, ~ 0.158 (14).

DISCUSSION

FETAL/JUVENILE MUTATION

Early cancer models posited that a constant number of adult cells were at continuous risk of acquisition of a required set of oncomutations that resulted in lethal tumors (15–17). Platt's questions about selective growth advantage of early oncomutants (18) stimulated the creation of the two-stage model of Armitage and Doll in which initiation was restricted to a constant number of adult cells (1). An alternate theory, fetal/juvenile or, more narrowly, gestational initiation was based on Luria's demonstration that in a bacterial culture growing from a small number of cells early

mutations give rise to large mutant colonies and late mutations to small ones (10). This concept was clearly applicable to the stem cells of a growing organ. By maturity a stem cell initiated early in development would create a large preneoplastic colony, an initiation in the last juvenile doubling would create a single initiated stem cell. Early cancers would arise from early developmental initiation, cancers of extreme old age would arise from the last juvenile initiation events (30, 31, 39, 40). In the algebraic model provided here each successive net doublings of stem cells of the fetal/juvenile period provides initiated colonies. Deaths from colonies initiated in any specific doubling interval would be distributed over subsequent years of life until all persons with initiated colonies have died (Figure 2).

The Armitage–Doll two-stage model, $CAL(h,t)$, of age-specific cancer rates amended to incorporate a (a) defined maximum number of pre-maturation colonic epithelial stem cells and (b) limitation of tumor initiating mutations to the fetal/juvenile period accurately predicts the observed incidence function $INC(h,t)$ for colorectal cancer in the cohort examined.

Biological data supporting limitation of mutations to the fetal/juvenile period were scant but suggestive. Brash and Ponten reported that increases particular point mutant *TP53* colonies in human skin were restricted to the juvenile years and that subsequent solar exposure in adults increased mutant colony size

but not number (35, 36). More recently, the distributions of five nuclear and 17 mitochondrial point mutations in adult human lung epithelium were found to match a simple Luria–Delbruck expansion of mutant stem cells for ten stem cell doublings prior to maturity. Both the nuclear and mitochondrial mutant fractions assayed in human lung epithelium remained constant over the age range of patients, 35–76 years (6).

Consistent with the fetal/juvenile limitation of tumor initiating mutations are the findings that (a) sunlight does not induce new nuclear mutations in adult skin (35, 36) and (b) cigarette smoke does not induce new nuclear or mitochondrial mutations in adult human lung epithelium (6, 9). There appear to be no data indicating that mutant fractions increase in adult solid organs (41) although increases with age in circulating leukocytes have been reported, see (6) for review. This requires rethinking about human carcinogenesis in which it is generally believed (and taught) that environmental mutagens continuously initiate tumors by acting on adult stem cells despite the absence of any supportive data.

MUTATOR/HYPERMUTABLE METAKARYOTIC STEM CELLS OF ORGANOGENESIS AND CARCINOGENESIS

The observed high fetal/juvenile mutation rates have been associated with amitotic, non-eukaryotic cells that arise in the fourth-to fifth-week of gestation that appear to serve as the stem cells of human organogenesis and carcinogenesis (2, 3, 6, 32). These large “metakaryotic” mononuclear cells or syncytia with bell shaped nuclei are found in all human fetal/juvenile organs, preneoplastic lesions, and neoplasias. They increase by symmetric amitotic fission and have been observed to produce all parenchymal, subsequently mitotic, cell forms by asymmetrical mitotic fissions in tissues derived from all three primordial germ layers (4). Metakaryotes also display bizarre modes of chromatin organization (5) and of DNA segregation that occurs prior to, or concomitant with, DNA replication in sister cells (4). Curiously, the replication process appears to involve DNA copying by the error-prone DNA polymerase β insofar as about half of cancer-initiating *APC* point mutation hotspots sampled are attributable to errors of DNA polymerase β copying undamaged DNA *in vitro* (42).

NUMBER (n) AND RATES OF REQUIRED INITIATING MUTATIONS

Within the algebraic expression of $CAL(h, t)$ the term $(\Pi R_i)^{1/n}$ represents the geometrical mean of “ n ” initiating mutations. Estimates of this parameter derived from best fits of the model given the assumption of population stratification for $n = 2$ and $m = 1$ –5 were 3.5, 2.3, 2.8, 2.8, and 4.3×10^{-5} respectively (Figure 3) agreeing well with estimates from clinical enumeration of *APC*^{-/-} mutant polyps and *OAT* gene-inactivating mutation rates of ~ 2 – 4×10^{-5} per stem cell doubling during colon organogenesis (11, 12). For $n = 1$, $m = 1$, $F < 1$, the rate estimate of $(\Pi R_i)^{1/n}$ is about 5.3×10^{-8} , for $n = 3$, $m = 1$, $F < 1$ about 2.8×10^{-4} . Values of the geometrical mean of initiation mutations assuming values of n other than 2 are thus clearly discordant with *APC* and *OAT* colonic mutation rate estimates. These facts, derived from clinical genetic observations in both inherited and sporadic forms of colorectal cancers are wholly consistent with the conclusion that $n = 2$ and inconsistent with values of $n \neq 2$.

NUMBER (m) AND RATES OF REQUIRED PROMOTION MUTATIONS

It is not mathematically possible to estimate “ m ” (Tables 3, 4) or the related geometrical mean of promotion mutation rates, $(\Pi R_A)^{1/m}$ by simply fitting $CAL(h, t)$ to $INC(h, t)$.

For $n = 2$ and $m = 1$, and $F \leq 1$, we note that the estimated values of a hypothetical promotion mutation, $(\Pi R_A)^{1/m}$, range from 1 to 5×10^{-5} , approximating the estimated geometric mean rate of 2 – 5×10^{-5} calculated for gene initiating mutations in *APC* or *OAT* genes in the fetal/juvenile stem cells (11, 12).

For $n = 2$ and $m = 4$ or 5 and $F \leq 1$ the values of $(\Pi R_A)^{1/m}$ bracket the estimated mean rate of losses of heterozygosity (LOH) of polymorphic markers in human adenocarcinomas of $\sim 8 \times 10^{-3}$. This calculation depended on the assumption that LOH in colonic adenocarcinomas arose from events in preneoplastic stem cells and not for events in post promotion neoplastic stem cell divisions. This conundrum points to the need for measurements of point and larger chromosomal mutations in human colonic adenomas.

While many kinds of genetic or epigenetic events might be required in preneoplastic metakaryotic stem cells, the data suggest that oncogene activation mutations, limited to a small set of amino acid substitutions in a proto-oncogene, is an unlikely candidate. Such events are expected to occur at rates some 100 time lower than gene-inactivation rates via point mutations (43). Gene deletion pathways as represented by loss of heterozygosity of polymorphic markers could not result in oncogene activation. To date no example of a required proto-oncogene activation has been found for any major form of human cancer (6).

However, the possibility of a large number of independent promotion pathways each with lower event rates than calculated here cannot be excluded (problem of potential parallel promotion pathways). In a formal sense we cannot even exclude $m = 0$ because no genetic or other rare required event has been identified for promotion for any human tumor.

ESTIMATION OF PRENEOPLASTIC GROWTH RATE, μ

Using the best-fit stipulated values for $n = 2$, $F < 1$ and any value of $m = 1$ –5, the preneoplastic growth rate “ μ ” was estimated to range from ~ 0.192 to 0.156 . This may be compared to the juvenile growth rate of mass in males, 0.158 , and females, 0.163 (14) but somewhat lower than Armitage and Doll’s original estimate that death rates doubled about every 5 years, i.e., a doubling rate of 0.2 (1). For the conditions $n = 2$, $m = 4$ or 5 our estimates of μ in preneoplastic colonic adenomas brackets the estimate of male juvenile growth rates. It would seem prudent to consider the possibility that preneoplastic stem cell growth rates might even more closely approximate the juvenile growth rates than previously suggested (14).

ENVIRONMENTAL CANCER RISKS DURING THE FETAL/JUVENILE PERIOD

Agreement between the origins of adult somatic mutations in mutator/hypermutable fetal/juvenile human stem cells and age-specific cancer rates offers an explanation of epidemiological associations between fetal and early childhood exposure to known mutagenic stimuli, particularly sunlight (35).

Generational changes of age and organ-specific cancer rates in immigrant populations toward those of the new country of residence may also be thought of in terms of fetal/juvenile initiation mutations. Adult immigrants would have experienced tumor initiation in the country of origin while children conceived and/or raised in the new country would experience mutagenic stimuli differing from those of their parents' land. In both older and younger immigrants the environment of the new country would define promotional stresses that might act in adults by inducing promotional oncomutations and/or by selecting through growth stimulation conditionally initiated stem cells acquired in the fetal/juvenile period (44).

Similarly, the marked decrease in death rates from lower G.I. tract (**Figures 1C,D**) and other incurable cancers in children and young adults post 1940 may reasonably be attributed to a decline in fetal/juvenile initiation mutation rates that began circa 1940 and continued through 2006 (see text footnote 1). We note in passing that these changes began very soon after vitamin supplementation was begun in the U.S., but the constant exponential decreases persisting for more than 60 years may suggest other hypotheses.

Save for sunlight exposure of juveniles, however, there is no direct evidence that environmental mutagens cause oncomutations in humans (41). "Spontaneous" mutation caused by simple DNA polymerase misincorporation errors or errors following DNA damage by endogenous processes could account for all fetal/juvenile mutations save for juvenile mutation by sunlight. Environmental mutagens may have been or still may be indirectly responsible for some of these oncomutations. For these reasons we cannot distinguish between the hypotheses that metakaryotic stem cells of organogenesis are constitutively mutator (high spontaneous mutation rate), hypermutable (markedly susceptible to endogenous or exogenous mutagenic agents), or both.

THE ASSUMPTION OF POPULATION HOMOGENEITY

When it was posited that the population is in some way inhomogeneous with regard to risk, $F < 1$, the fit of the model, $CAL(h,t)$ to the recorded $INC(h,t)$ was significantly improved (**Figure 3**). One form of risk stratification is represented by the \sim tenfold range in colonic adenomatous polyp numbers observed among individuals of FAP families (11) that suggest a range of fetal/juvenile *APC* mutation rates. A similar variation of numbers of mutant crypts displaying loss of the second allele of (*OAT*) has been observed in adult colon (12). The remaining discrepancies between the two-stage model adjusted for fetal/juvenile initiation and the colorectal cancer incidence data may thus lie principally in the assumption of a homogeneous oncomutation rate that assumes a Poisson distribution of initiating events among persons in each organogenic stem cell doubling period. This oversimplification is compensated only in part by positing a value of $F < 1$ (**Table 4**, **Figure 3**).

THE ASSUMPTION OF SYNCHRONOUS COMPETING FORMS OF DEATH

Assuming a competing synchronous form of death with shared risk(s) with colorectal cancer does not further increase goodness of fit under best-fit conditions (**Figure 3**; **Table 3**) in accord with epidemiological studies of familial colorectal cancer (34). This possibility must however be considered for cancers such as breast

and ovarian cancers in which synchronous age-specific death rates display shared familial risk.

THE ASSUMPTION OF A SINGLE PATHWAY OF INITIATION AND PROMOTION

We and others modeling carcinogenesis have so far treated a special limited case in which there is a single or predominant pathway to a lethal cancer. Thus, for instance in the case of colorectal cancer we assume a single pathway of initiation involving mutations i and j that appear to be independent losses of the maternal and paternal alleles of the *APC* gene conjoined to a single pathway of promotion mutations A, B, C, D, ... However, it is already known that a minor pathway to colon cancer exists involving point mutations in the beta-catenin gene (45) so our calculations for initiation rates are in fact a weighted average of the *APC*, beta-catenin, and other unknown minor pathways.

A general case would encompass multiple initiation pathways linked to one or more promotional pathways are responsible for cancers in organs such as the lungs, prostate, breast, and pancreas for which single tumor suppressor genes have not been found.

SUMMARY

These quantitative analyses/ruminations point to an amended qualitative model of human carcinogenesis at least for colorectal carcinogenesis. It seems that as metakaryotic stem cells grow to form the colon they experience a very high mutation rate and that in each organogenic stem cell doubling newly initiated stem cells are created. The newly initiated cells of each stem cell doubling grow at the same rate as their sister uninitiated stem cells during the fetal/juvenile period but after maturity grow at a constant rate similar to the juvenile growth rate of about 0.16 per year. Any initiated stem cell may experience additional promotion mutations sufficient to create a first neoplastic stem cell that gives rise to a rapidly growing tumor (adenocarcinoma) that may kill *in situ* or through metastases.

If for any organogenic stem cell all necessary initiation and promotion mutations occur before maturity a pediatric tumor would be expected. But after maturity only stem cell lineages initiated before maturity are expected to grow and create the exponential increase in cancer deaths with adult age. Eventually, the subpopulation that has experienced colorectal stem cell initiation is depleted in the surviving population in extreme old age and death rates from this form of cancer in the surviving population decline.

When matched to fetal/juvenile *APC* mutation rates derived from observation of polyp number in FAP patients the number of initiation mutations, n , is consistent with two and only two initiation events. A value of $m = 4$ or 5 is suggested if it is assumed that gene-inactivating events occurring at rates estimate from LOH fractions in adenocarcinomas are required for promotion. Population risk stratification for initiation mutation rates is indicated by a wide distribution of mutant numbers in adult colons and by the fact accounting in part for said stratification within the amended two-stage model improves the concordance with age-specific colorectal cancer rates. These findings point to the importance of mutagenesis in fetal and juvenile stem cells as primary determinants of adult age-specific cancer rates and suggest that they account for inter-generational

shifts in cancer risks among organs in immigrant populations and the observed exponential decline of recent decades in pediatric cancer mortality in solid organs such as the colon (Figure 1D).

ACKNOWLEDGMENTS

We gratefully acknowledge the contributions of the late Professor Lars Ehrenberg (Wallenberg Laboratory and the University of

Stockholm) and those of Dr. Suresh Moolgavkar (Fred Hutchinson Cancer Research Center, Seattle, WA, USA). Jonathan Jackson and Aaron Fernandes contributed to the initial design of Cancer-Fit. Five of us (Lohith G. Kini, Tushar Kamath, Rebecca Kusko, Elena V. Gostjeva, William G. Thilly) were supported in part by a research contract to M.I.T. from United Therapeutics Corp., Silver Spring, MD, USA for the study of the role of metakaryotic cells in human carcinogenesis.

REFERENCES

- Armitage P, Doll R. A two-stage theory of carcinogenesis in relation to the age distribution of human cancer. *Br J Cancer* (1957) **92**:161–9. doi:10.1038/bjc.1957.22
- Gostjeva EV, Thilly WG. Stem cell stages and the origins of colon cancer: a multi-disciplinary perspective. *Stem Cell Rev* (2005) **1**:243–52. doi:10.1385/SCR:1:3:243
- Gostjeva EV, Zukerberg L, Chung D, Thilly WG. Bell-shaped nuclei dividing by symmetrical and asymmetrical nuclear fission have qualities of stem cells in human colonic embryogenesis and carcinogenesis. *Cancer Genet Cytogenet* (2006) **164**:16–24. doi:10.1016/j.cancergencyto.2005.05.005
- Gostjeva EV, Koledova V, Tomita-Mitchell A, Mitchell M, Goetsch MA, Varmuza S, et al. Metakaryotic stem cell lineages in organogenesis of humans and other metazoans. *Organogenesis* (2009) **5**:191–200. doi:10.4161/org.5.4.9632
- Gruhl AN, Gostjeva EV, Thilly WG, Fomina JN, Darroudi F. Human fetal/tumor metakaryotic stem cells: pangenomic homologous pairing and telomeric end-joining of chromatids. *Cancer Genet Cytogenet* (2010) **203**:203–8. doi:10.1016/j.cancergencyto.2010.08.015
- Sudo H, Li-Sucholeiki X-C, Marcelino LA, Gruhl AN, Herrero-Jimenez P, Zarbl H, et al. Fetal-juvenile origins of point mutations in the adult tracheal-bronchial epithelium: absence of detectable effects of age, gender or smoking status. *Mutat Res* (2008) **646**:25–40. doi:10.1016/j.mrfmmm.2008.08.016
- Cohnheim J. *Vorlesungen ueber allgemeine Pathologie. Ein Handbuch fuer Aertzte und Studierende*. Berlin: Hirschwald (1877).
- Boveri T. *The Origin of Malignant Tumors*. [English Translation by M. Boveri (1929) Baillière, Tindall & Cox]. Baltimore: The Williams and Wilkins Company (1914).
- Coller HA, Khrapko K, Torres A, Frampton MW, Utell MJ, Thilly WG. Mutational spectra of a 100-base pair mitochondrial DNA target sequence in bronchial epithelial cells: a comparison of smoking and nonsmoking twins. *Cancer Res* (1998) **58**:1268–77.
- Luria SE, Delbruck M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* (1943) **28**:491–511.
- Kwak EL, Chung DC. Hereditary colon cancer syndromes: an overview. *Clin Colorectal Cancer* (2007) **6**:340–4. doi:10.3816/CCC.2007.n.002
- Campbell F, Williams GT, Appleton MA, Dixon MF, Harris M, Williams ED. Post-irradiation somatic mutation and clonal stabilisation time in the human colon. *Gut* (1996) **39**:569–73. doi:10.1136/gut.39.4.569
- Penman BW, Crespi CL, Komives EA, Liber HL, Thilly WG. Mutation of human lymphoblasts exposed to low concentrations of chemical mutagens for long periods of time. *Mutat Res* (1983) **108**:417–36. doi:10.1016/0027-5107(83)90137-9
- Herrero-Jimenez P, Tomita-Mitchell A, Furth EE, Morgenthaler S, Thilly WG. Population risk and physiological parameters for colon cancer. The union of an explicit model for carcinogenesis with the public health records of the United States. *Mutat Res* (2000) **447**:73–116. doi:10.1016/S0027-5107(99)00201-8
- Holloman JH, Fisher JC. Nucleation and growth of cell colonies. *Science* (1950) **111**:489–91. doi:10.1126/science.111.2888.489
- Nordling CO. A new theory on the cancer-inducing mechanism. *Br J Cancer* (1953) **7**:68–72. doi:10.1038/bjc.1953.8
- Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br J Cancer* (1954) **81**:1–12. doi:10.1038/bjc.1954.1
- Platt R. Clonal aging and cancer. *Lancet* (1955) **265**:867. doi:10.1016/S0140-6736(55)90453-0
- Friedwald WF, Rous P. The initiating and promoting elements in tumor production, An analysis of the effects of tar, benzpyrene and methylcholanthrene on rabbit skin. *J Exp Med* (1944) **80**:101–26. doi:10.1084/jem.80.2.101
- Herrero-Jimenez P, Thilly G, Southam PJ, Tomita-Mitchell A, Morgenthaler S, Furth EE, et al. Mutation, cell kinetics, and subpopulations at risk for colon cancer in the United States. *Mutat Res* (1998) **400**:553–78. doi:10.1016/S0027-5107(98)00067-0
- Cook PJ, Doll R, Fillingham SA. A mathematical model for the age distribution of cancer in man. *Int J Cancer* (1969) **41**:93–112. doi:10.1002/ijc.2910040113
- Harding C, Pompei F, Wilson R. Cancer suppression at old age. *Cancer Res* (2008) **68**:4465–78. doi:10.1158/0008-5472.CAN-07-1670
- Harding C, Pompei F, Wilson R. Peak and decline in cancer incidence, mortality and prevalence at old ages. *Cancer* (2011) **118**:1371–86. doi:10.1002/cncr.26376
- Knudson AG. Two genetic hits (more or less) to cancer. *Nat Rev Cancer* (2001) **1**:157–62. doi:10.1038/35101031
- Knudson AG. Mutations and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* (1971) **68**:820–3. doi:10.1073/pnas.68.4.820
- Moolgavkar SH, Knudson AG Jr. Mutation and cancer: a model for human carcinogenesis. *J Natl Cancer Inst* (1981) **66**:1037–52.
- Beerenwinkel N, Antal T, Dingli D, Traulsen A, Kinzler KW, Vogelstein B, et al. Genetic progression and the waiting time to cancer. *PLoS Comput Biol* (2007) **3**:e225. doi:10.1371/journal.pcbi.0030225
- Salk JJ, Fox EJ, Loeb LA. Mutational heterogeneity in human cancers: origin and consequences. *Annu Rev Pathol* (2010) **5**:51–75. doi:10.1146/annurev-pathol-121808-102113
- Frank SA. Somatic mosaicism and cancer: inference based on a conditional Luria-Delbrück distribution. *J Theor Biol* (2003) **223**:405–12. doi:10.1016/S0022-5193(03)00117-6
- Moolgavkar SH. Biologically motivated two-stage model for cancer risk assessment. *Toxicol Lett* (1988) **43**:139–50.
- Thilly WG. Looking ahead: algebraic thinking about genetics, cell kinetics and cancer. In: Bartsch H, Hemminki K, O'Neill IK editors. *Methods for Detecting DNA Damaging Agents in Humans: Applications in Cancer Epidemiology and Prevention*. IARC Monograph No. 89. Lyon: IARC Scientific Publications (1988). p. 486–92.
- Sudo H, Li-Sucholeiki XC, Marcelino LA, Gruhl AN, Zarbl H, Willey JC, et al. Distributions of five common point mutants in the human tracheal-bronchial epithelium. *Mutat Res* (2006) **596**:113–27. doi:10.1016/j.mrfmmm.2005.12.008
- Peto R. Epidemiology, multistage models, and short-term mutagenicity tests. In: Hiatt HH, Winsten JA, editors. *Origins of Human Cancer*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory (1977). p. 1403–29.
- Hemminki K, Granstrom C, Chen B. The Swedish family-cancer database: update, application to colorectal cancer and clinical relevance. *Hered Cancer Clin Pract* (2005) **3**:7–18. doi:10.1186/1897-4287-3-1-7
- Brash DE, Rudolph JA, Simon JA, Lin A, McKenna GJ, Baden HP, et al. A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. *Proc Natl Acad Sci USA* (1991) **88**:10124–8. doi:10.1073/pnas.88.22.10124
- Brash DE, Ponten J. Skin precancer. *Cancer Surv* (1998) **32**:69–113.
- Berge T, Lundberg S. Cancer in Malmo 1958–1969. An autopsy study. *Acta Pathol Microbiol Scand Suppl* (1977) **260**:1–235.
- Linstrom P, Janzon L, Sternby NH. Declining autopsy rate in Sweden: a study of causes and consequences in Malmo. *Sweden J Intern Med* (1997) **242**:157–65. doi:10.1046/j.1365-2796.1997.00178.x
- Thilly WG. What actually causes cancer? *Technol Rev* (1991):48–54.
- Meza R, Luebeck EG, Moolgavkar SH. Gestational mutations and carcinogenesis. *Math Biosci* (2005) **197**:188–210. doi:10.1016/j.mbs.2005.06.003

41. Thilly WG. Have environmental mutagens caused oncomutations in people? *Nat Genet* (2003) **34**:255–9. doi:10.1038/ng1205
 42. Muniappan BP, Thilly WG. The DNA polymerase beta replication error spectrum in the adenomatous polyposis coli gene contains human colon tumor mutational hotspots. *Cancer Res* (2002) **62**:3271–5.
 43. Leong PM, Thilly WG, Morgenthaler S. Variance estimation in single-cell assays: comparison to experimental observations in human lymphoblasts at 4 gene loci. *Mutat Res* (1985) **150**:403–10. doi:10.1016/0027-5107(85)90137-X
 44. Hemminki K, Ji J, Brandt A, Mousavi SM, Sundquist J. The Swedish Family-Cancer Database 2009: prospects for histology-specific and immigrant studies. *Int J Cancer* (2010) **126**:2259–67. doi:10.1002/ijc.24795
 45. Morin PJ, Sparks AB, Korinek V, Barker N, Clevers H, Vogelstein B, et al. Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science* (1997) **275**:1787–90. doi:10.1126/science.275.5307.1787
- Conflict of Interest Statement:** Our work has been supported in part by a best-effort research contract from United Therapeutics Corp. of Silver Springs, MD, USA to study metakaryotic cells in carcinogenesis. This support is listed in the acknowledgments. No payments to any co-author or expectation of patent royalties attends upon this theoretical exploration of human colorectal cancer.
- Received: 08 May 2013; accepted: 07 October 2013; published online: 29 October 2013.*
- Citation:* Kini LG, Herrero-Jimenez P, Kamath T, Sanghvi J, Gutierrez E Jr, Hensle D, Kogel J, Kusko R, Rexer K, Kurzweil R, Refinetti P, Morgenthaler S, Koledova VV, Gostjeva EV and Thilly WG (2013) Mutator/hypermutable fetal/juvenile metakaryotic stem cells and human colorectal carcinogenesis. *Front. Oncol.* **3**:267. doi: 10.3389/fonc.2013.00267
- This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology.*
- Copyright © 2013 Kini, Herrero-Jimenez, Kamath, Sanghvi, Gutierrez, Hensle, Kogel, Kusko, Rexer, Kurzweil, Refinetti, Morgenthaler, Koledova, Gostjeva and Thilly. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Left-right symmetry breaking in mice by left-right dynein may occur via a biased chromatid segregation mechanism, without directly involving the *Nodal* gene

Stephan Sauer* and Amar J. S. Klar*

Gene Regulation and Chromosome Biology Laboratory, Frederick National Laboratory for Cancer Research, Frederick, MD, USA

Edited by:

James L. Sherley, Boston Biomedical Research Institute, USA

Reviewed by:

Paola Parrella, IRCCS Casa Sollievo Della Sofferenza, Italy
Mitsuru Furusawa, Neo-Morgan Laboratory Incorporated, Japan

*Correspondence:

Stephan Sauer and Amar J. S. Klar,
Gene Regulation and Chromosome Biology Laboratory, Frederick National Lab for Cancer Research, 7th Street Fort Detrick, Frederick, MD 21702, USA.
e-mail: sauers@mail.nih.gov;
klara@mail.nih.gov

Ever since cloning the classic *iv* (*inversed viscerum*) mutation identified the “left-right dynein” (*Ird*) gene in mice, most research on body laterality determination has focused on its function in motile cilia at the node embryonic organizer. This model is attractive, as it links chirality of cilia architecture to asymmetry development. However, *Ird* is also expressed in blastocysts and embryonic stem cells, where it was shown to bias the segregation of recombined sister chromatids away from each other in mitosis. These data suggested that *Ird* is part of a cellular mechanism that recognizes and selectively segregates sister chromatids based on their replication history: old “Watson” versus old “Crick” strands. We previously proposed that the mouse left-right axis is established via an asymmetric cell division prior to/or during gastrulation. In this model, left-right dynein selectively segregates epigenetically differentiated sister chromatids harboring a hypothetical “left-right axis development 1” (“*Ira1*”) gene during the left-right axis establishing cell division. Here, asymmetry development would be ultimately governed by the chirality of the cytoskeleton and the DNA molecule. Our model predicts that randomization of chromatid segregation in *Ird* mutants should produce embryos with 25% situs solitus, 25% situs inversus, and 50% embryonic death due to heterotaxia and isomerism. Here we confirmed this prediction by using two distinct *Ird* mutant alleles. Other than *Ird*, thus far *Nodal* gene is the most upstream function implicated in visceral organs laterality determination. We next tested whether the *Nodal* gene constitutes the *Ira1* gene hypothesized in the model by testing mutant’s effect on 50% embryonic lethality observed in *Ird* mutants. Since *Nodal* mutation did not suppress lethality, we conclude that *Nodal* is not equivalent to the *Ira1* gene. In summary, we describe the origin of 50% lethality in *Ird* mutant mice not yet explained by any other laterality-generating hypothesis.

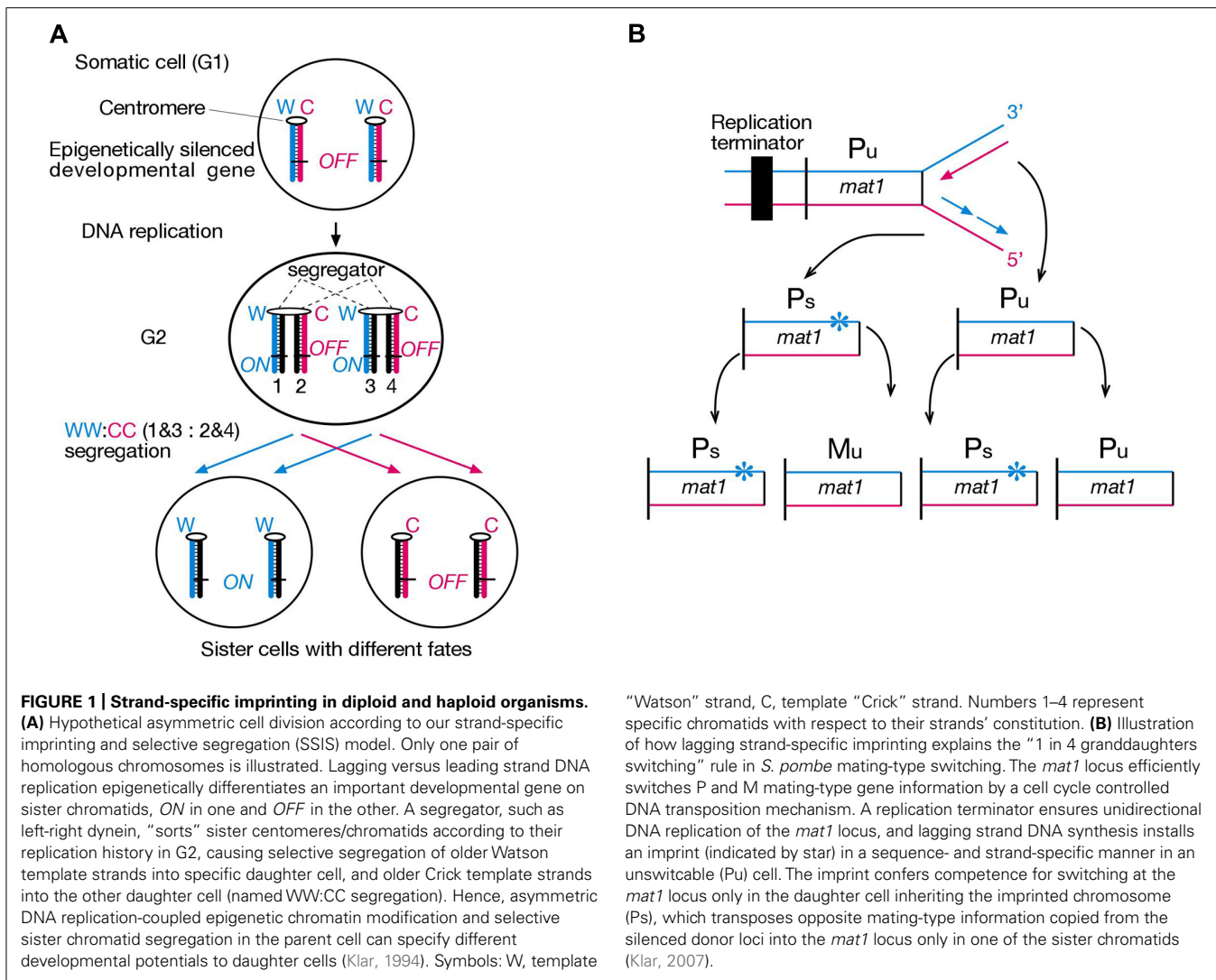
Keywords: laterality development, left-right dynein, asymmetric cell division, DNA strands differentiation, selective chromatid segregation

INTRODUCTION

It is crucial for multicellular development that cells possess a memory system, which ensures stable inheritance of acquired developmental states during development of tissues and organs of an organism. The field of epigenetics studies this cellular memory system, and “epigenetic” is often defined as “mitotically heritable changes in gene expression that do not involve modulation of the primary DNA sequence.” For development, it is equally important that cells are able to change their acquired developmental state and differentiate along evolutionarily defined lineage paths. A crucial question is how epigenetic information can be changed and passed onto developmentally differentiated sister cells during asymmetric cell division. We proposed a solution to this problem. Namely, sister chromatids can be epigenetically differentiated regarding a developmentally important gene during S-Phase, based on lagging versus leading strand DNA replication, followed by selective sister chromatid segregation to specific daughter cells (**Figure 1A**). Our Somatic Strand-specific Imprinting and selective sister chromatid Segregation (SSIS) model (Klar, 1994) postulates that a specific daughter inherits both template Watson and first time

synthesized Crick strand-containing (WC’) homologous chromosomes, thereby the other daughter inherits with both new Watson and old Crick (W’C) homologous chromosomes (referred to as WW:CC segregation pattern). As a consequence, a single gene or a gene cluster is poised for expression in one daughter cell and silenced in the other daughter cell. Likewise, if sister chromatids were selectively segregated in a WC:WC fashion, then both daughter cells would inherit equivalent epigenetic make ups and hence retain similar developmental potentials, as seen in symmetrical stem cell divisions. The SSIS model is based on studies on fission yeast (*Schizosaccharomyces pombe*) mating-type switching (Klar, 2007), and has been tested *in vitro* in mouse embryonic stem (ES) cells (Armakolas and Klar, 2006, 2007), and *in vivo* in a mouse model for body laterality development (this study).

Schizosaccharomyces pombe is a haploid unicellular eukaryote, whose cells either express P or M mating-type information from the alternate alleles of the *mat1* locus residing in chromosome 2. The *mat1* mating-type content switches between M and P information by a cell cycle controlled DNA transposition



mechanism, such that one out of four granddaughter cells switches cell type and expresses the mating-type opposite to that of the grandmother cell (Figure 1B). Genetic and biochemical analysis revealed that mating-type switching is controlled by lagging-versus leading-strand DNA replication at the *mat1* locus. In particular, lagging-strand DNA synthesis installs an imprint at *mat1* (most probably a two nucleotide long DNA:RNA hybrid from an incompletely removed Okazaki fragment), which initiates a double-strand break during the following S-Phase to start the DNA transposition event that underlies *mat1* switching. Hence developmental asymmetry between sister cells can be traced back to double helical structure of the *mat1* gene and lagging- versus leading-strand synthesis of specific DNA strands in two consecutive cell divisions (Klar, 2007).

We proposed that a similar mechanism might produce asymmetric cell divisions in diploid organisms by epigenetic means as well. First, strand-specific imprinting would epigenetically differentiate sister chromatids in S-Phase, and selective segregation of thus differentiated sister chromatids would create sister cells with different developmental fates. This model is called SSIS, and

was initially developed by us to explain internal organ laterality development in vertebrates (Klar, 1994).

The development of bilateral asymmetry can be conceptually divided into three steps: First comes the initial symmetry-breaking event, usually ascribed to cellular amplification of a molecular chirality. This is followed by differential gene expression in cell fields on either side of the midline, which translates to step three, left/right (L/R) asymmetric organogenesis (Aw and Levin, 2009). For internal organ situs development in vertebrates, a great deal of molecular understanding has been achieved to decipher steps two and three, where many molecular pathways, seemingly conserved between model organisms, have been defined and well accepted (Nakamura and Hamada, 2012). For example, the TGF- β related signaling molecule Nodal is conserved in all deuterostomes examined, and usually specifies the left body-side (Chea et al., 2005). Its activity is inhibited toward the midline by Nodal's own transcriptional targets of the Lefty family of diffusible molecules, which represents a prime example of a reaction-diffusion mechanism (Nakamura et al., 2006; Muller et al., 2012). In contrast, identity of the symmetry-breaking event, the "first event," that initiates

left-biased Nodal expression is controversial, because no unifying mechanism between vertebrate phyla has been identified to date (Vandenberg and Levin, 2009). Some vertebrates such as mice, frogs, and zebrafish are proposed to employ motile cilia during gastrulation at equivalent embryonic organizer regions, known as the node, gastrocoel roof plate, and Kupffer vesicle, respectively. Cilia's beating is thought to either transport a morphogen leftwards in extraembryonic space (Nonaka et al., 1998), or induce asymmetrical calcium signaling in conjunction with mechanosensory cilia (McGrath et al., 2003). As a consequence, Nodal signaling is induced more strongly in left-sided neighboring tissues (lateral plate mesoderm in the mouse), and its autoregulatory feedback loop with Lefty molecules confers robustness to the signaling cascade (Nakamura and Hamada, 2012). This model is very attractive as it links the molecular chirality of the cilium and its building blocks to chirality of the developing embryo. However, several observations prominently question this model's universality, and some data would rather support a role for nodal cilia during step two of bilateral asymmetry development, namely, asymmetric gene expression on either side of the midline. First, pigs, for example, undergo L/R axis development without motile nodal cilia, undermining a universal role for motile cilia in vertebrate and mammalian symmetry-breaking (Vandenberg and Levin, 2010). Second, in species that employ cilia, a number of genes that are required for proper nodal cilia motility and positioning are also expressed in non-ciliated cells at much earlier embryonic stages. Examples include planar cell polarity genes *Vangl2* and *Dvl2*, *inversin* and *left-right dynein* (Aw and Levin, 2009). Thus, it is unclear whether these proteins exert their critical function in L/R axis development at the node. Third, mouse blastomere cells rearrangement has been shown to influence direction of embryonic turning, indicating that some aspects of laterality development certainly occur prior to gastrulation, and are independent of nodal cilia (Gardner, 2010). Last, both zebrafish and mouse mutants have been isolated, which show Kupffer vesicle or node ciliary defects but no L/R phenotypes, and *vice versa* (Vandenberg and Levin, 2010). Therefore, despite overwhelming evidence suggesting that cilia do have an important function in L/R asymmetry development in several species, they are unlikely to truly control initial symmetry-breaking in the embryo to generate L/R asymmetry (Tabin, 2005; Klar, 2008; Lobikin et al., 2012).

In 1959, Hummel and Chapman (1959) first described the recessive *iv* (*inversed viscerum*) mutation, where 50% of homozygous mice develop situs inversus (i.e., mirror-image reversal of internal organs), and 50% have normal organ situs. Parental organ situs does not affect organ situs of the offspring, thus this mutation randomizes L/R asymmetry. More detailed analysis revealed high rates of heterotaxia (random and independent sidedness of internal organs) affecting both normal and situs inversus homozygous mutants at similar rates and severity (Layton, 1978). This suggests that in addition to its involvement in the first step of asymmetry development, the *iv* gene product is also needed in the second and/or third conceptual steps described above. Molecular cloning by Supp et al. (1997) showed that the *iv* mutation changed a highly conserved glutamic acid to lysine within the motor domain of a dynein heavy chain gene, which was thereafter

named *left-right dynein* (*lrd*). *Lrd* message was detected in blastocysts and (blastocyst-derived) ES cells, ventral node cells, and some ciliated embryonic and adult epithelia. It was classified as an axonemal dynein despite its obvious expression in many non-ciliated cell types. At the time the authors were not aware that node cells contain motile cilia, and even concluded that: "... embryonic expression indicates that mechanisms other than ciliary movement are involved in L/R specification" (Supp et al., 1997). Later it was found that Node cells contained ciliated cells (Nonaka et al., 1998) whose motility is dependent on *lrd* (Supp et al., 1999). Technically difficult studies further showed that beating nodal cilia created a leftward fluid-flow in extraembryonic space (Nonaka et al., 1998), and artificial fluid-flow reversal had a dominant effect on situs development in normal and *lrd* mutant embryos (Nonaka et al., 2002). These data clearly highlight the node's function as an embryonic organizer during L/R axis development. Whether the L/R asymmetry is truly established by nodal flow or whether this simply represents a "back-up" mechanism remains to be determined. To address this question, our lab has started to generate a conditional allele for *lrd* to discriminate between early cytoplasmic and later axonemal (cilia) functions.

A study from our lab has provided genetic evidence that *lrd* does indeed have a functional role in non-ciliated cells (Armakolas and Klar, 2007). Liu et al. (2002) had engineered mouse ES cell lines, which allowed for selection of Cre/loxP-mediated mitotic recombinants between homologs. If recombination happens in G2, recombined chromatids can either segregate together (Z segregation) or into different sister cells (X segregation). X segregants thereby acquire homozygosity of any heterozygous marker distal to the crossover site. Interestingly, centromere-proximal loxP sites on chromosome 7 (DT1E9 cell line) always led to X segregation, whereas loxP sites on chromosome 11 or further centromere-distal on chromosome 7, produced the usually expected random mix of X and Z segregants (Liu et al., 2002). Differentiation of DT1E9 ES cells to endoderm cells preserved the exclusive X segregation pattern, whereas neuroectoderm cells showed exclusive Z segregation. Three other *in vitro* differentiated cell types showed random patterns (Armakolas and Klar, 2006). We proposed that cell type-specific biased segregation patterns were due to selective chromatid recombination as well as selective segregation of chromosome 7 sister chromatids in mitosis. Remarkably, *lrd* mRNA expression was evident in ES, endoderm, and neuroectoderm cells, and RNAi-mediated knockdown randomized segregation patterns, consistent with our SSIS model (Klar, 2008). In this model, *lrd* would "sort" sister chromatids based on their replication history and selectively segregate sister centromeres to sister cells (Figure 1A). We propose that *lrd*'s function in non-ciliated cells is to bias sister chromatid segregation of one or a specific set of chromosomes. By theory, this function is not confined to a single L/R axis establishing asymmetric cell division, but probably happens in other developmental contexts where asymmetric or strictly symmetric cell divisions occur. Additional support for this is provided by *lrd*'s expression profile available on the gene annotation portal biogps.org, where *lrd* shows high expression in hematopoietic stem cells (<http://www.biogps.org/#goto>). Here we tested developmental biology predictions of the SSIS model concerning the *lrd* mutant. In a second experiment we tested whether

the *Nodal* gene comprises the “left-right axis development 1” (*lra1*) gene specified in the SSIS model.

MATERIALS AND METHODS

MOUSE BREEDING AND HUSBANDRY

Lrd-Neo-GFP mice were a kind gift from Dr. Martina Brueckner at Yale University, New Haven, CT. The iv stock (EM:02531) was purchased (live) from EMMA repository, Harwell, UK. Delta *Nodal* mice were a kind gift from Dr. Michael Kuehn, Frederick National Laboratory, MD. All mice were kept according to Animal Care and User Committee (ACUC) guidelines, Frederick National Laboratory, MD.

GENOTYPING

Between 3 and 4 weeks of age, tailclips were performed according to ACUC guidelines. Tails were digested by overnight incubation at 55°C in 200 µl of tail buffer [100 mM NaCl, 10 mM Tris-HCl pH 7.5, 10 mM EDTA, 0.5% (w/v) *N*-Lauroylsarcosine, 100 µg/ml Proteinase K]. The solution was then diluted 1:1 with dH₂O, 1 µl was used for PCR reactions. *Lrd*-Neo-GFP primers: wtaF3: CTCTGCAGGCAGAGCGGCT, taR3: GCTTGCCGGTGGTGCAGA, wtR3: CGGGTCTAGGGCAAAGCGTT. PCR: 95°C 2 min – 34× (94°C 20 s, –64.5°C 20 s, –72°C 30 s) 72°C 5 min. wt allele: 194 bp, targeted allele: 266 bp. *Nodal* Delta primers: F4299: CAGAAGAGGGATTGGGGTTTGCAG, R4457: GATCGGAAGTCTAGGAACCTAGAAAC. 95°C 2 min – 32× (94°C 30 s, –65°C 30 s, –72°C 30 s) 72°C 5 min. Targeted (delta) allele: ~180 bp. iv primers: 1959 TaqI F: GCTAACCACCAACCACATGCTG, 1959 TaqI R: CACGGATTCCAGCCCAGATC. 25 µl PCR product was digested with 25 U of Taq alpha I (NEB) in a 40 µl reaction, at 65°C for 45 min. The iv mutation destroys the Taq alpha I site in the PCR fragment. wt bands: 92 bp, iv band: 184 bp.

RESULTS

A TEST OF A KEY PREDICTION OF THE SSIS MODEL

Our model makes several testable predictions for the phenotype of the *lrd* mouse mutant. First, randomization of sister chromatid segregation during the critical L/R axis establishing cell division should have three different outcomes: 25% WW:CC cell pairs leading to normal organ situs later in development, 25% CC:WW cell pairs leading to inversed organ situs, and 50% WC:WC cell pairs causing embryonic lethality or death soon after birth due to isomerism (mirror-image sidedness of organs) or heterotaxia (random and independent sidedness of organs; **Figure 2**). Lethality occurs because of the *lra1* gene's ON/OFF epiallele constitution in both sister cells. Prediction of 50% lethality in *lrd* mutant mice is a major difference between SSIS hypothesis and mainstream nodal cilia hypotheses for L/R axis development (Klar, 2008).

We acquired two different *lrd* mutant mouse strains, the original *iv* strain from EMMA repository and the *Lrd*-Neo-GFP mouse from Dr. Martina Brueckner's laboratory (McGrath et al., 2003). The *iv* strain originated from a complex mixed background until siblings were inbred for >20 generations (EMMA repository, personal communication). The *Lrd*-Neo-GFP allele was introduced into ES cells of Sv129 genetic background (McGrath et al., 2003). To reduce background specific influences, both

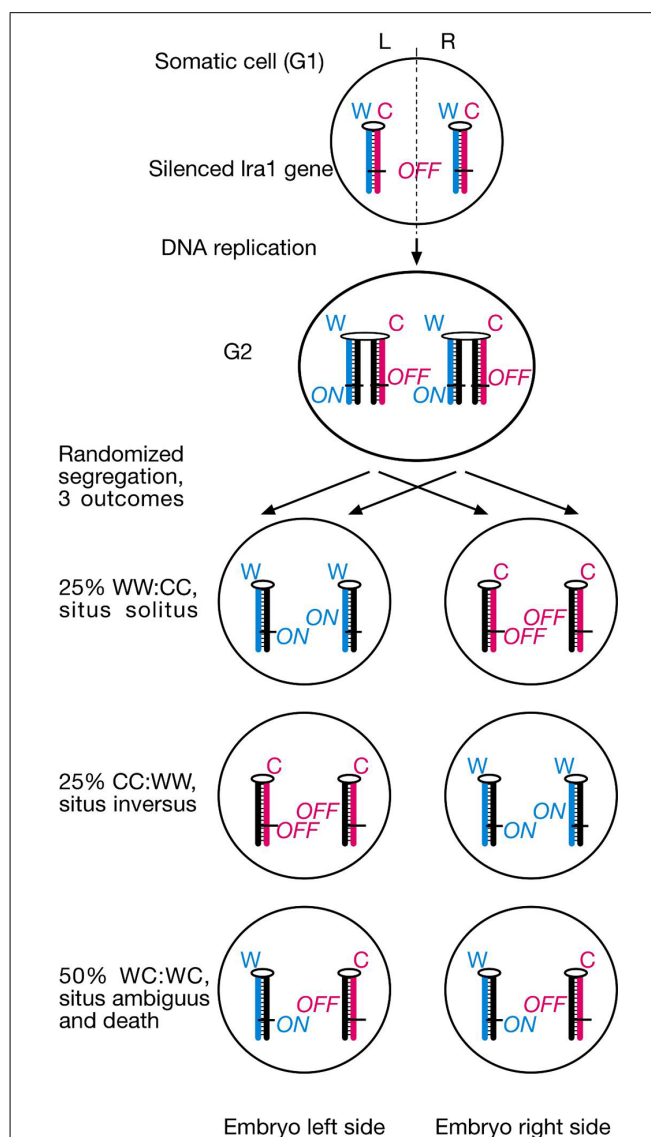


FIGURE 2 | SSIS-predictions concerning embryo situs and survival rates of *lrd* mutants. Proposed laterality-generating asymmetric cell division is randomized in the *lrd* mutant. The future L/R axis is set by cytoplasmic polarization and alignment of a single cell with respect to the anterior-posterior and dorsal-ventral body axes. Sister chromatids containing a hypothetical “leftness-encoding” left-right axis-establishing gene 1 (*lra1*) are epigenetically differentiated. Normally, left-right dynein would selectively segregate older Watson template strand-containing sister chromatids harboring *lra1* “ON” epialleles to the left body side, and older Crick template strand-containing sister chromatids harboring *lra1* “OFF” epialleles to the right body side as described in Figure 1A. Randomized segregation due to left-right dynein mutation will result in three different outcomes shown here: 25% WW:CC cell pairs, causing normal situs development, 25% CC:WW cell pairs, causing development of situs inversus, and 50% WC:WC cell pairs, causing severe developmental situs abnormalities incompatible with survival.

strains were bred onto C57BL/6 strain for one generation. Lethality rates were determined at weaning age (3–4 weeks) by PCR based genotyping of tailclip DNA (see section Materials and Methods).

Table 1 | Observed rates of allele frequencies: *Lrd-Neo-GFP* allele, *lrd*^{+/-} x *lrd*^{+/-}.

<i>Lrd</i> ^{+/+} (%)	<i>lrd</i> ^{+/-} (%)	<i>lrd</i> ^{-/-} (%)
53 (32)	90 (55)	22 (13)

Table 2 | Observed rates of allele frequencies: *iv* allele, *iv*^{+/-} x *iv*^{-/-}.

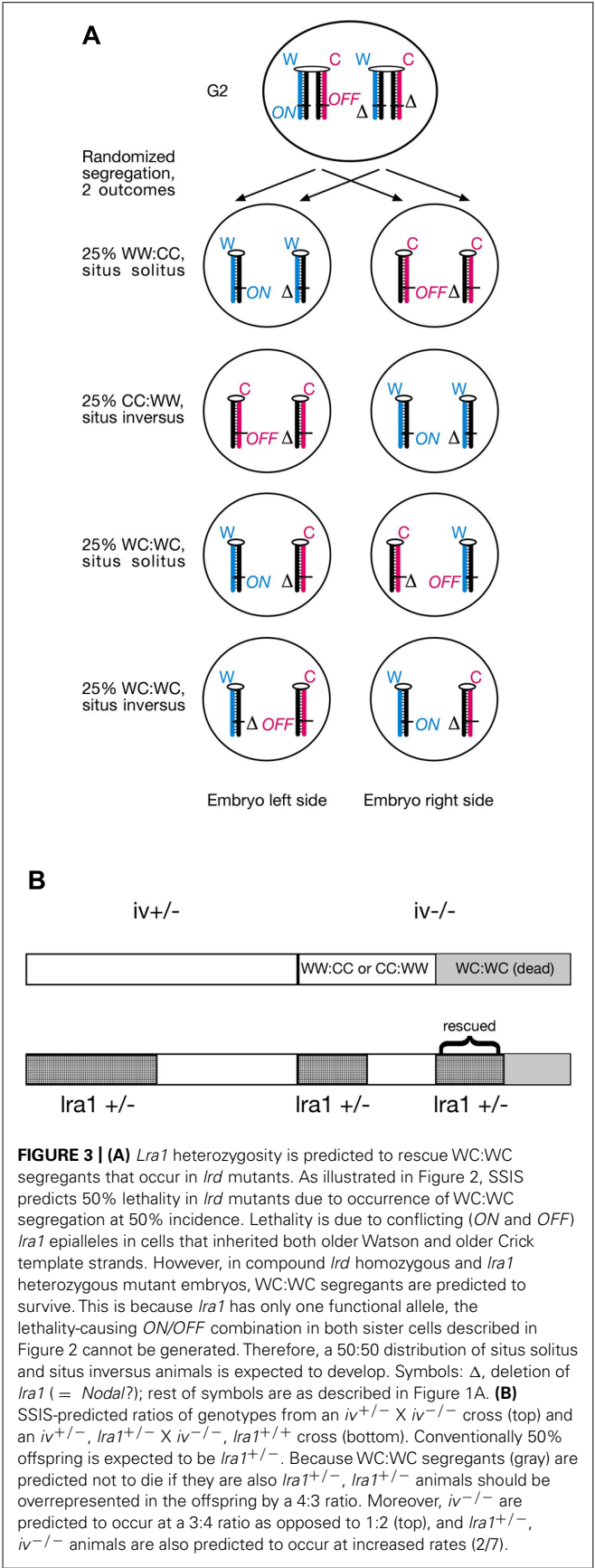
<i>iv</i> ^{+/-} (%)	<i>iv</i> ^{-/-} (%)
74 (67)	37 (33)

Lrd-Neo-GFP mice carry a *GFP-lrd exon 1* fusion as well as a *Neo* cassette on the opposite strand of *lrd intron 1*. Since the *Neo* transgene is under the control of a very strong promoter and transcribed antisense to *lrd*, *lrd* transcription is effectively shut down and homozygous mutant mice are indistinguishable from true knock-out mice: 50% of live animals exhibit situs inversus (McGrath et al., 2003). Several heterozygous intercrosses (*lrd*^{+/-}) were set up and DNA from tails from 165 offspring was analyzed (Table 1). We detected 53 *lrd*^{+/+} : 90 *lrd*^{+/-} : 22 *lrd*^{-/-} animals. The SSIS hypothesis predicts ~24 (165/7) of live-born mice to be *lrd*^{-/-}. This is because 1/8 (half of 1/4 animals with ^{-/-} genotype) of the initial number of homozygous mutant mice is expected to live, 1/8 is expected to die and thus reduce the total number of mice that are available for analysis to 7/8. As a result, 1/8 of the initial mice correspond to 1/7 of observable mice. If lethality was not an issue, then ~41 mice (1/4 of 165) should have the *lrd*^{-/-} genotype. Our observed number of 22 *lrd*^{-/-} mice is statistically equivalent to the SSIS-predicted number of 23.57 (*p*-value of ~0.6, chi-square test).

Encouraged by the heterozygous cross results, we set up four *iv*^{+/-} X *iv*^{-/-} crosses. The results are summarized in Table 2. Conventionally 1/2 of the offspring is expected to be *iv*^{-/-}. However if lethality affected 50% of the *iv*^{-/-} mice, this fraction would be reduced to 1/3 among live animals. Analysis of 111 offspring revealed 74 *iv*^{+/-} and 37 *iv*^{-/-} mice, which meets SSIS prediction exactly.

DOES NODAL CONSTITUTE THE *LRA1* GENE HYPOTHESIZED IN THE SSIS MODEL?

According to the SSIS model, heterozygosity for the *lra1* gene would prevent embryonic lethality in *iv*^{-/-} embryos because heterotaxia or isomerism would not occur. Consequently, 50% would develop normal organ situs and 50% would develop situs inversus in embryos with *lra1*^{+/-}, *iv*^{-/-} genotype (Figure 3A). We chose a candidate gene approach, and considered the *Nodal* gene as a likely candidate for the *lra1* gene as it is the gene, other than *iv*, that functions most upstream in the L/R pathway. *Nodal* belongs to the TGF-β family of extracellular signaling molecules and has been shown to be amongst the earliest asymmetrically (left-sided) expressed molecules in a variety of species, ranging from snails to man (Nakamura and Hamada, 2012). Since *Nodal* is essential for mesoderm induction during gastrulation and the



mutant embryos die before L/R axis is established (Lowe et al., 2001), its function for L/R axis development is difficult to address by using a conventional null allele. Conditional inactivation of the FoxH1 transcriptional activator in the lateral plate mesoderm causes loss of *Nodal* expression and R/R isomerism (Yamamoto et al., 2003). Likewise, injection of *Nodal*^{-/-} ES cells into wild-type (wt) blastocysts results in development of R/R isomers (Oh and Li, 2002). Thus, *Nodal* is considered to encode “leftness.” If *Nodal* is in fact the *lra1* gene, the 50% lethality of *lrd* homozygous mutant embryos should be suppressed in *Nodal*^{+/-} heterozygotes according to our model (Figure 3A).

We determined whether heterozygosity for a null allele of the *Nodal* gene (Lowe et al., 2001) in *iv* crosses affected lethality ratios. Specifically, we determined whether the mutation suppresses 50% lethality of *iv*^{-/-} mice described in Figure 2. Because the *Nodal* gene deletion is homozygous lethal, we therefore quantitated viability of only heterozygous animals. We generated several males heterozygous for *iv* and *delta Nodal* mutations, which we set up with *iv*^{-/-}, *Nodal*^{+/+} females. Therefore, half of all offspring will be heterozygous for *delta Nodal* allele. In this mating set up, several predictions concerning ratios of expected genotypes are made (Figure 3B). Should heterozygous *Nodal* mutation not influence lethality ratios, then 1/3 *iv*^{-/-} mice should be observed, just like the result of the cross described in Table 2. Accordingly, 50% will be heterozygous for the *delta Nodal* mutation, and 1/6 (1/3 × 1/2) will be both *iv*^{-/-} and carriers of *delta Nodal*. If heterozygosity for *Nodal* rescues lethality in WC:WC segregants, then only 1/8 (1/2 × 1/2 × 1/2) of initial conceptuses will die, which reduces the total number of observable mice to 7/8. As discussed above, 1/8 of initial mice correspond to 1/7 of observable (live) mice. Three out of seven of live mice will be of *iv*^{-/-} genotype, and 4/7 will be *Nodal*^{+/-}, should this mutation suppress lethality in a subgroup of mice destined to die. Moreover, the ratio of *iv*^{-/-} and *Nodal*^{+/-} animals will now be not a simple product of their individual ratios, rather this genotype will be enriched, and is predicted to occur at a 2/7 rate: 1/7 stems from WW:CC (or CC:WW) segregants and 1/7 from rescued WC:WC segregants (Figure 3B). The observed result of these crosses is summarized in Table 3. Amongst 202 offspring, we found 66 *iv*^{-/-} and 103 *Nodal*^{+/-} animals. Thirty-three mice were both *iv*^{-/-} and *Nodal*^{+/-}. These numbers do not support the *Nodal* gene being the hypothetical *lra1* gene. Rather, they show that *lrd* mutation causes 50% lethality in *Nodal* heterozygotes as well.

DISCUSSION

We propose DNA’s chirality and its asymmetric mode of replication as a potential source for installing binary imprints on the

chromatin fiber, and selective segregation of thus differentiated sister chromatids to sister cells as a novel and largely uncharacterized molecular mechanism associated with asymmetric cell divisions. The *lrd*-dependent segregation bias of mouse chromosome 7 sister chromatids in mitotic recombination experiments involving ES cells, endoderm cells, and neuroectoderm (Armakolas and Klar, 2007) cells could represent a case for selective sister chromatid segregation. Even though direct evidence for this interpretation is still missing, it led us to further investigate the phenotype of the *lrd* mouse mutant. In our model *lrd* functions to “sort” and selectively segregate sister chromatids based on their replication history in a WW:CC fashion. This L/R symmetry-breaking asymmetric cell division would be oriented along the L/R axis, positional information for it would presumably come from polarized cytoskeleton (Vandenberg and Levin, 2009). Randomization of chromatid segregation in *iv*^{-/-} mice would lead to 25% normal organ situs (WW:CC segregants), 25% situs inversus (CC:WW segregants), and 50% death (WC:WC segregants). The 50 situs solitus : 50 situs inversus distribution in *lrd* mutant live animals has been described in numerous studies, therefore we only focused on assessment of lethality ratios by studying Mendelian inheritance of *lrd* mutant alleles in appropriate genetic crosses. In order to eliminate potential allele-specific or genetic background-specific artifacts, we analyzed two distinct *lrd* null alleles that had been outbred onto mixed backgrounds. Both crosses revealed *lrd* homozygous mutant animals at rates 50% below Mendelian predictions. This result is consistent with our SSIS hypothesis even though it does not provide definitive proof of it. Nearly all studies of mouse laterality stress only the 50% situs solitus: 50% situs inversus phenotype of *iv*^{-/-} mice and ignore the 50% lethality phenotype. Approximately 50% lethality was first noted by Layton (1978) in one of the earliest studies of *iv* mutant crosses. Our results presented here with *iv* confirmed the estimations of Layton (1978) and extended it to the newly made *Lrd-Neo-GFP* allele. One caveat for our SSIS explanation is that the original *iv* allele might be a leaky missense mutation generating the observed effects. It was therefore important to investigate phenotypes of a different allele, which is why we used the second *Lrd-Neo-GFP* allele for our analysis. A third allele was investigated previously, but the analysis was very limited to draw conclusions regarding lethality (Supp et al., 1999). Unfortunately, that allele was not saved (M. Bruckner, personal communication).

We next sought to test another prediction of our model, namely that heterozygosity for the hypothetical *lra1* gene would rescue WC:WC segregants. The rationale therefore is that strand-specific imprinting of *lra1* would lead to conflicting (ON/OFF) *lra1* epialleles in both WC:WC sister cells. If one allele of *lra1* is a null allele (due to heterozygosity), then different epialleles cannot be conflicting anymore (Figure 3A). We chose a reverse genetics approach and tested the *Nodal* gene as a possible candidate for *lra1*. Analysis of >200 offspring did not show a protective function for *Nodal* heterozygosity in *lrd* mutant animals: therefore, *Nodal* cannot be *lra1*. We did however confirm the 50% lethality phenotype of *iv*^{-/-} genotype, indicating that lethality was not affected by *Nodal* gene dosage.

We have eliminated *Nodal* as a candidate for *lra1*, and its ActR2B receptor can also be disregarded, because a study from

Table 3 | Allele frequencies in offspring of *iv*^{+/-} *Nodal*^{+/-} × *iv*^{-/-} *Nodal*^{+/-} cross.

n = 202	Conv. expected	SSIS expected	Observed
<i>iv</i> ^{-/-}	1/3 = 67.3	3/7 = 86.6	66
<i>Nodal</i> ^{+/-}	1/2 = 101	4/7 = 115.4	103
<i>iv</i> ^{-/-} and <i>Nodal</i> ^{+/-}	1/6 = 33.7	2/7 = 57.7	32

En Li's laboratory (Oh and Li, 2002) recorded situs ambiguous (pulmonary isomerism) in around 40% of *iv*^{-/-} *ActR2B*^{+/-} embryos. If *ActR2B* were *lral*, then the SSIS model predicts occurrence of situs solitus and situs inversus only. The *Nodal* signaling pathway is highly complex and regulated by numerous factors on several levels (Schier, 2009). After *Nodal* precursor is activated by proprotein convertases and released into extracellular space, Lefty proteins limit its activity via a reaction-diffusion mechanism. *Nodal* binding to activin receptors is assisted by distinct co-receptors, and sometimes *Nodal* binds in conjunction with other TGF- β molecules as a heterodimer. Moreover, evidence from zebrafish suggests tight post-transcriptional control of *Nodal*, Lefty and activin receptor gene expression by microRNAs (Schier, 2009). Given this level of complexity, a reverse genetic screen is unfeasible to tackle the identity of *lral* gene in the first step of L/R asymmetry development.

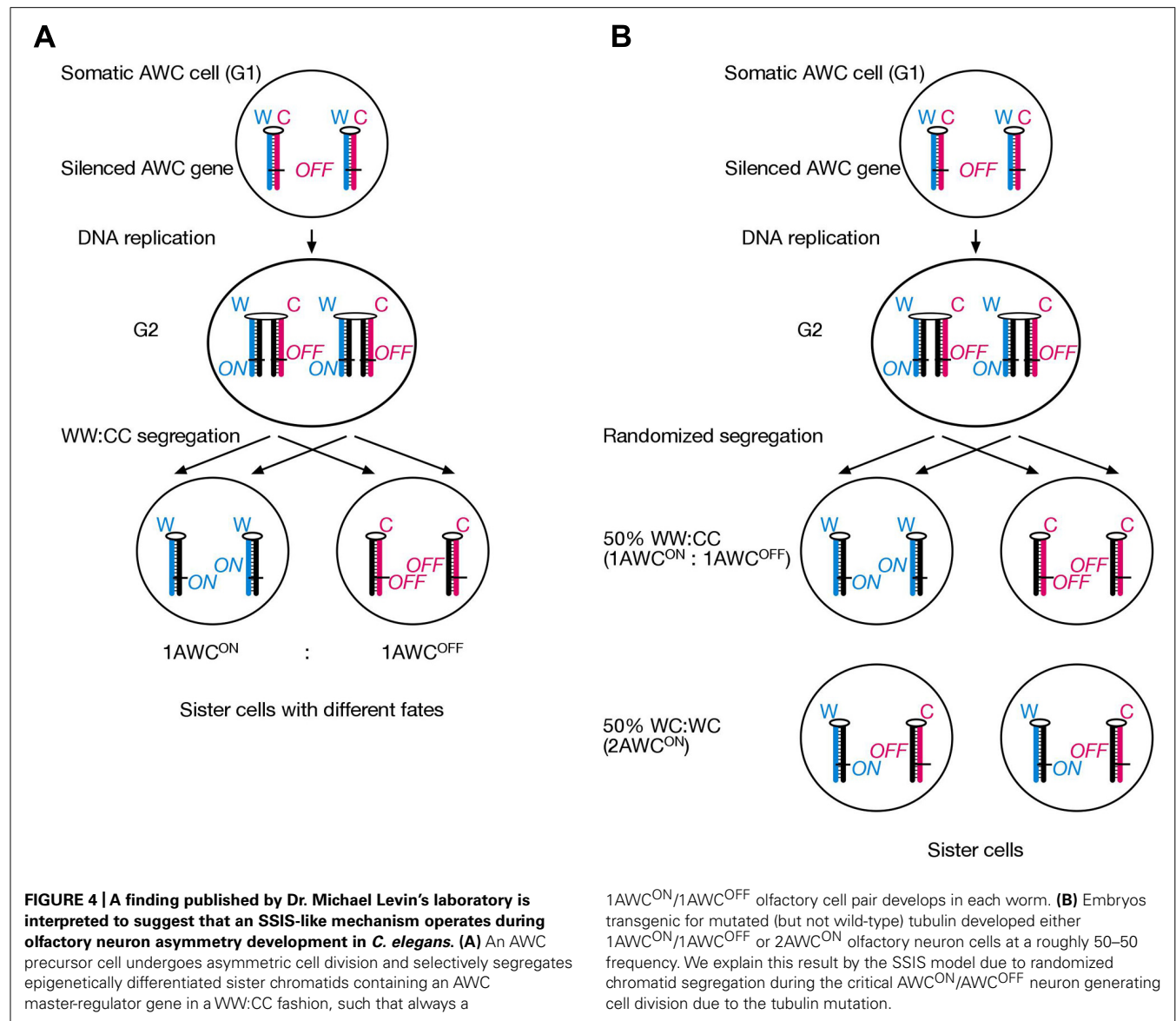
Interestingly, two recent studies have suggested that the nematode *C. elegans* employs SSIS mechanism during Neuronal asymmetry development. A study from Michael Levin's research group provides genetic support that an SSIS-type asymmetric cell division operates in olfactory neuron development, although the evidence has not been interpreted as such by the authors. The Levin laboratory has a long-standing interest in vertebrate L/R axis development, and has highlighted the role of the cytoskeleton in cellular polarization for years (Aw and Levin, 2009; Vandenberg and Levin, 2009, 2010). It had come to the authors' attention that *Arabidopsis* mutants affecting radial flower symmetry were mapped to alpha-tubulin and a gamma-tubulin associated protein (Lobikin et al., 2012). Remarkably, introducing the same alpha-tubulin mutation into *Xenopus* 1-cell embryos resulted in development of heterotaxia, and in cultured human HL-60 cells it disturbed the leftward bias (with respect to the nucleus-centrosome axis) of pseudopodia protrusion. In addition, *C. elegans* "AWC" olfactory neural asymmetry was also affected by mutating a tubulin homolog (TBA-9, 75% amino acid identity with *Arabidopsis* alpha-tubulin) at two conserved amino acids. In wt worms, the AWC neuron is in the "ON" state (AWC^{ON}) on one body side and in the "OFF" state (AWC^{OFF}) on the other body side; sidedness is stochastic (Chang et al., 2011). This developmental asymmetry can be visualized by introducing the "str-2p::GFP" fluorescent GFP construct in the genome. Importantly, chromosomal integration site of the GFP transgene is irrelevant for faithful AWC^{ON} versus AWC^{OFF} discrimination, indicating that the cause for asymmetric GFP expression acts in trans for the transgene. The authors chose this model system for body asymmetry development studies, because the AWC^{ON} and AWC^{OFF} cells show cytoskeletal polarization and asymmetric calcium signaling, which is sensitive to the microtubule depolymerizing drug nocodazole (Chang et al., 2011). Overexpression of wt TBA-9 tubulin in transgenic worms causes only mild laterality defects, with 82% of worms displaying the normal 1AWC^{ON}/1AWC^{OFF} phenotype. Overexpression of mutant TBA-9, in contrast, results in 42% normal 1AWC^{ON}/1AWC^{OFF} and 45% novel 2AWC^{ON} "heterotaxic" phenotype. This roughly 50:50 distribution is consistent with a SSIS mechanism operating in the mother AWC cell (Figure 4). We hypothesize that one daughter inherits normally two AWC^{ON} epialleles and the other two AWC^{OFF} epialleles are inherited by the

other daughter cell (WW:CC segregation). Unlike our SSIS model for mouse L/R axis development, this asymmetric distribution of sister chromatids in the worm occurs irrespective of the L/R body axis. We propose that introduction of mutated tubulin renders the AWC cell's cytoskeleton unable to direct selective chromatid segregation in mitosis, hence a novel WC:WC segregation results at 50% frequency. A simple explanation for 2AWC^{ON} phenotype in WC:WC segregants would be dominance of the AWC^{ON} over the AWC^{OFF} epiallele.

A second study implicating an SSIS-like asymmetric cell division in *C. elegans* neuronal asymmetry development has been recently published by Horvitz/Stillman laboratories (Nakano et al., 2011). Here, a GFP-reporter screen served to isolate mutants that changed the paired asymmetric MI motor neuron/e3D epidermal cell pair to a symmetrical e3D cell pair on both sides of the brain. Positional cloning identified a gain-of-function mutation in a *histone H3* gene that deleted its last 11 amino acids, thereby impairing the ability to form H3/H4 tetramers during chromatin assembly. Likewise, RNAi against chromatin assembly factor 1 (CAF-1) or PCNA pheno-copied the H3 mutant. The authors suggest that newly lagging strand synthesized DNA contains elevated levels of PCNA and associated CAF-1 containing histone chaperone complex, which deposits higher nucleosome density. This could represent an epigenetic imprint in itself, or serve to nucleate covalent chromatin modifications. The latter seems somewhat more likely, since the epigenetic imprint is transmitted through several mitoses, as it is the MI/e3D great-great-grandmother cell that directs development of distinct cell fates three cell divisions later on. Like in our SSIS model, selective segregation of epigenetically differentiated sister chromatids is an integral part of the authors' model. However, neither direct nor indirect evidence is presented. Because genetic evidence suggests that mutated tubulin (Lobikin et al., 2012) randomizes (the normally selective) chromatid segregation during an AWC^{ON}/AWC^{OFF} olfactory neuron asymmetry generating cell division, we propose to test whether mutated TBA-9 also affects the MI/e3D neuronal asymmetry.

The SSIS model is conceptually based on three aspects: (i) differential chromatin imprinting during inherently lagging versus leading strand replication, (ii) one or several genes whose expression is affected by this imprint, and (iii) a segregator that identifies and "sorts" epigenetically differentiated sister chromatids by operating at sister centromeres in mitosis. We have presented genetic evidence for (iii), namely that *lrd* acts as a segregator in a L/R axis defining cell division in mouse. In contrast, Nakano et al. (2011) have provided evidence for (i), but have not identified the segregator. If the segregator can be identified, *C. elegans* will be excellently suited to use forward genetics to identify the gene or set of genes (ii) that are imprinted and selectively segregated, as outlined in our *iv*^{-/-}, *Nodal*^{+/-} breeding experiment.

We have highlighted two studies that support a SSIS-type mechanism in the development of neuronal asymmetries in *C. elegans*. Based on genetics of psychosis development in human carriers of balanced chromosome 11 translocations, we have previously proposed that a similar mechanism may operate during human brain lateralization (Klar, 2004). Analogous to our model for body laterality development, brain laterality development would also initiate with a single critical asymmetric cell division, where chromosome



11 sister chromatids are selectively segregated in a WW:CC fashion. If one chromosome 11 homolog is fused to the centromere of another chromosome not undergoing selective mitotic segregation, then WW:CC and WC:WC segregation for chromosome 11 are expected to occur at equal frequencies. Fifty per cent incidence of psychosis development in four different families with balanced chromosome 11 translocations support our hypothesis (Singh and Klar, 2007).

Whether asymmetric cell divisions elsewhere during normal tissue homeostasis employ a SSIS mechanism remains to be determined. If they do exist, then somatic chromosomal translocations could potentially randomize these asymmetric cell divisions and initiate tumorigenesis. An example would be a resting tissue stem cell that only enters the cell cycle upon tissue injury. It asymmetrically divides to produce a rapidly-proliferating transiently amplifying stem cell. This cellular asymmetry development would be controlled by asymmetric segregation of cytoplasmic

determinants, but also by WW:CC segregation of epigenetically differentiated sister chromatids, where cell cycle promoting genes remain silenced in the mother cell, but poised for expression in the transiently amplifying daughter cell. A chromosomal translocation involving the chromosome undergoing selective segregation in the tissue stem cell could therefore change the WW:CC pattern to a WC:WC pattern. As a result, the resting tissue stem cell would acquire proliferative capacities of the transiently amplifying stem cell, leading to neoplasia. Additional oncogenic mutations will eventually render this cell growth cancerous. Despite this example being rather simplistic, it should be appreciated that genes controlling asymmetric cell division are increasingly recognized as tumor suppressors. *Drosophila* *brat* and *prospero* mutants, for example, fail to undergo asymmetric neuroblast cell divisions, and develop larval brain tumors (Betschinger et al., 2006). We suggest that somatic chromosomal translocations in tissue stem cells could affect biased segregation of sister chromatids,

and change strictly asymmetrically dividing stem cells to stem cells that undergo symmetrical cell divisions in terms of epigenetic imprints on differentiated sister chromatids distal to the translocation breakpoint.

Curiously, a 1992 study published in *The Lancet* (Sandson et al., 1992) found a correlation of aberrant brain laterality development and breast cancer. Right-handed breast cancer patients and healthy controls were subjected to computer tomographic brain scans. Eighty-two per cent of control subjects showed left hemispheric dominance, whereas in the breast cancer group this number was reduced to 51%. Although this study should be cautiously interpreted until replicated elsewhere, it certainly suggests that brain laterality- and breast cancer-development share a common genetic pathway (Klar, 2011). We suggest that this pathway controls asymmetric cell divisions during embryonic brain development, and during cell turnover in the lactiferous duct upon periodic hormonal growth stimulation. Hence, improving our understanding of vertebrate laterality development could eventually impact on cancer prevention and treatment.

Taken together, 50% lethality phenotype in the *lrd* mouse mutant supports predictions made by the SSIS model for laterality development. Here, *lrd* is part of a cellular mechanism that selectively segregates epigenetically differentiated sister chromatids concerning their replication history with respect to a cytoskeleton-based early L/R axis (Klar, 2008; Vandenberg and Levin, 2010; Lobikin et al., 2012). The overwhelming majority of studies on *lrd* in the mouse have focused on its role in conferring nodal cilia motility. This is understandable, since genetics of spontaneous and targeted mouse mutations affecting laterality development have generally pointed to a central role for motile nodal cilia. Moreover, the earliest known molecular L/R asymmetries appear after node formation in the mouse. In chicken and *Xenopus*, in contrast, earlier asymmetries involving Gap-junctional communication (Levin and Mercola, 1999), H⁺/K⁺ ATPase activity (Levin et al., 2002), and serotonin signaling (Fukumoto et al., 2005) have been identified. As many of the studies on earlier asymmetry determinants in frogs and chicken involved embryo-exposure to pharmacological inhibitors, mouse embryo-culture protocols will need to vastly improve until replication can even be considered. In this regard it is noteworthy that one of the leading laboratories for mouse embryo *in vitro* culture has recently tested the relationship of nodal cilia emitted force and asymmetry development in

several mouse mutants affecting cilia biogenesis and motility (Shinohara et al., 2012). Surprisingly, the authors found that as few as two motile nodal cilia were sufficient to break bilateral symmetry. These data are rather compatible with the “2-cilia hypothesis,” which was initially postulated by the Hirokawa, Brueckner, and Tabin laboratories (Okada et al., 1999; McGrath et al., 2003; Tabin and Vogan, 2003). Here, mechanical force exerted by beating nodal cilia is read out by mechanosensory cilia, which are associated with polycystin-2 (a calcium release channel) to induce left-sided calcium signaling. Hence loss of *polycystin-2* is predicted to ablate calcium and Nodal signaling altogether. Pennekamp et al. (2002) indeed found loss of *Nodal* expression in the majority of *polycystin-2* deficient embryos, however the *Nodal* downstream target *Pitx2* showed bilateral expression. Clearly how body laterality is initially developed, whether in visceral organs or in the nervous system, remains controversial thus far. Further work is needed to differentiate if any of the prevailing hypotheses can satisfactorily explain body laterality development. The SSIS model is simple to understand in that an asymmetric cell division constitutes the root cause of development. In this model, developmental decisions are made through particulate matter consisting of ON/OFF epigenetic states of gene expression of developmentally important gene(s). Thus, in addition to acting as genetic material, DNA strands can provide the basis for evolution, cancer and development (Furusawa, 2011).

ACKNOWLEDGMENTS

We are grateful to Dr. Martina Brueckner (Yale University, New Haven, CT) for sharing the *Lrd-Neo-GFP* allele and to Dr. Michael Kuehn (Frederick National Laboratory, MD) for sharing the *delta Nodal* allele. We would like to thank Dr. Mark Lewandoski for discussions and use of his laboratory facilities and Lisa Dodge for mouse husbandry. The Intramural Research Program of the National Institutes of Health, Frederick National Laboratory for Cancer Research supports our research. Stephan Sauer is recipient of a long-term Fellowship from the Human Frontier Science Program Organization (LT-000444/2009).

AUTHOR CONTRIBUTIONS

Amar J. S. Klar and Stephan Sauer designed experiments, Stephan Sauer carried out experiments and collected data, Stephan Sauer and Amar J. S. Klar interpreted the data and wrote the manuscript.

REFERENCES

- Armakolas, A., and Klar, A. J. (2006). Cell type regulates selective segregation of mouse chromosome 7 DNA strands in mitosis. *Science* 311, 1146–1149.
- Armakolas, A., and Klar, A. J. (2007). Left-right dynein motor implicated in selective chromatid segregation in mouse cells. *Science* 315, 100–101.
- Aw, S., and Levin, M. (2009). Is left-right asymmetry a form of planar cell polarity? *Development* 136, 355–366.
- Betschinger, J., Mechtler, K., and Knoblich, J. A. (2006). Asymmetric segregation of the tumor suppressor *brat* regulates self-renewal in *Drosophila* neural stem cells. *Cell* 124, 1241–1253.
- Chang, C., Hsieh, Y. W., Lesch, B. J., Bargmann, C. I., and Chuang, C. F. (2011). Microtubule-based localization of a synaptic calcium-signaling complex is required for left-right neuronal asymmetry in *C. elegans*. *Development* 138, 3509–3518.
- Chea, H. K., Wright, C. V., and Swalla, B. J. (2005). Nodal signaling and the evolution of deuterostome gastrulation. *Dev. Dyn.* 234, 269–278.
- Fukumoto, T., Kema, I. P., and Levin, M. (2005). Serotonin signaling is a very early step in patterning of the left-right axis in chick and frog embryos. *Curr. Biol.* 15, 794–803.
- Furusawa, M. (2011). Implications of double-stranded DNA structure for development, cancer and evolution. *Open J. Genet.* 1, 78–87.
- Gardner, R. L. (2010). Normal bias in the direction of fetal rotation depends on blastomere composition during early cleavage in the mouse. *PLoS ONE* 5, e9610. doi: 10.1371/journal.pone.0009610
- Hummel, K. P., and Chapman, D. B. (1959). Visceral inversion and associated anomalies in the mouse. *J. Hered.* 50, 9–13.
- Klar, A. J. (1994). A model for specification of the left-right axis in vertebrates. *Trends Genet.* 10, 392–396.
- Klar, A. J. (2004). An epigenetic hypothesis for human brain laterality, handedness, and psychosis development. *Cold Spring Harb. Symp. Quant. Biol.* 69, 499–506.
- Klar, A. J. (2011). Breast cancer predisposition and brain hemispheric laterality specification likely share a common genetic cause. *Breast Dis.* 33, 49–52.
- Klar, A. J. (2007). Lessons learned from studies of fission yeast mating-type switching and silencing. *Annu. Rev. Genet.* 41, 213–236.

- Klar, A. J. (2008). Support for the selective chromatid segregation hypothesis advanced for the mechanism of left-right body axis development in mice. *Breast Dis.* 29, 47–56.
- Layton, W. M. Jr. (1978). Heart malformations in mice homozygous for a gene causing situs inversus. *Birth Defects Orig. Artic. Ser.* 14, 277–293.
- Levin, M., and Mercola, M. (1999). Gap junction-mediated transfer of left-right patterning signals in the early chick blastoderm is upstream of Shh asymmetry in the node. *Development* 126, 4703–4714.
- Levin, M., Thorlin, T., Robinson, K. R., Nogi, T., and Mercola, M. (2002). Asymmetries in H^+/K^+ -ATPase and cell membrane potentials comprise a very early step in leftright patterning. *Cell* 111, 77–89.
- Liu, P., Jenkins, N. A., and Copeland, N. G. (2002). Efficient Cre-loxP-induced mitotic recombination in mouse embryonic stem cells. *Nat. Genet.* 30, 66–72.
- Lobikin, M., Wang, G., Xu, J., Hsieh, Y. W., Chuang, C. F., Lemire, J. M., et al. (2012). Early, nonciliary role for microtubule proteins in left-right patterning is conserved across kingdoms. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12586–12591.
- Lowe, L. A., Yamada, S., and Kuehn, M. R. (2001). Genetic dissection of nodal function in patterning the mouse embryo. *Development* 128, 1831–1843.
- McGrath, J., Somlo, S., Makova, S., Tian, X., and Brueckner, M. (2003). Two populations of node monocilia initiate left-right asymmetry in the mouse. *Cell* 114, 61–73.
- Muller, P., Rogers, K. W., Jordan, B. M., Lee, J. S., Robson, D., Ramanathan, S., et al. (2012). Differential diffusivity of Nodal and Lefty underlies a reaction-diffusion patterning system. *Science* 336, 721–724.
- Nakamura, T., and Hamada, H. (2012). Left-right patterning: conserved and divergent mechanisms. *Development* 139, 3257–3262.
- Nakamura, T., Mine, N., Nakaguchi, E., Mochizuki, A., Yamamoto, M., Yashiro, K., et al. (2006). Generation of robust left-right asymmetry in the mouse embryo requires a self-enhancement and lateral-inhibition system. *Dev. Cell* 11, 495–504.
- Nakano, S., Stillman, B., and Horvitz, H. R. (2011). Replication-coupled chromatin assembly generates a neuronal bilateral asymmetry in *C. elegans*. *Cell* 147, 1525–1536.
- Nonaka, S., Tanaka, Y., Okada, Y., Takeda, S., Harada, A., Kanai, Y., et al. (1998). Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. *Cell* 95, 829–837.
- Nonaka, S., Shiratori, H., Saijoh, Y., and Hamada, H. (2002). Determination of left-right patterning of the mouse embryo by artificial nodal flow. *Nature* 418, 96–99.
- Oh, S. P., and Li, E. (2002). Gene-dosage-sensitive genetic interactions between inversus viscerum (iv), nodal, and activin type IIB receptor (ActRIIB) genes in asymmetrical patterning of the visceral organs along the left-right axis. *Dev. Dyn.* 224, 279–290.
- Okada, Y., Nonaka, S., Tanaka, Y., Saijoh, Y., Hamada, H., and Hirokawa, N. (1999). Abnormal nodal flow precedes situs inversus in iv and inv mice. *Mol. Cell* 4, 459–468.
- Pennekamp, P., Karcher, C., Fischer, A., Schweickert, A., Skryabin, B., Horst, J., et al. (2002). The ion channel polycystin-2 is required for left-right axis determination in mice. *Curr. Biol.* 12, 938–943.
- Sandson, T. A., Wen, P. Y., and LeMay, M. (1992). Reversed cerebral asymmetry in women with breast cancer. *Lancet* 339, 523–524.
- Schier, A. F. (2009). Nodal morphogens. *Cold Spring Harb. Perspect. Biol.* 1, a003459.
- Shinohara, K., Kawasumi, A., Takamatsu, A., Yoshida, S., Botilde, Y., Motoyama, N., et al. (2012). Two rotating cilia in the node cavity are sufficient to break left-right symmetry in the mouse embryo. *Nat. Commun.* 3, 622.
- Singh, G., and Klar, A. J. (2007). A hypothesis for how chromosome 11 translocations cause psychiatric disorders. *Genetics* 177, 1259–1262.
- Supp, D. M., Brueckner, M., Kuehn, M. R., Witte, D. P., Lowe, L. A., McGrath, J., et al. (1999). Targeted deletion of the ATP binding domain of left-right dynein confirms its role in specifying development of left-right asymmetries. *Development* 126, 5495–5504.
- Supp, D. M., Witte, D. P., Potter, S. S., and Brueckner, M. (1997). Mutation of an axonemal dynein affects left-right asymmetry in inversus viscerum mice. *Nature* 389, 963–966.
- Tabin, C. (2005). Do we know anything about how left-right asymmetry is first established in the vertebrate embryo? *J. Mol. Histol.* 36, 317–323.
- Tabin, C. J., and Vogan, K. J. (2003). A two-cilia model for vertebrate left-right axis specification. *Genes Dev.* 17, 1–6.
- Vandenberg, L. N., and Levin, M. (2010). Far from solved: a perspective on what we know about early mechanisms of left-right asymmetry. *Dev. Dyn.* 239, 3131–3146.
- Vandenberg, L. N., and Levin, M. (2009). Perspectives and open problems in the early phases of left-right patterning. *Semin. Cell Dev. Biol.* 20, 456–463.
- Yamamoto, M., Mine, N., Mochida, K., Sakai, Y., Saijoh, Y., Meno, C., et al. (2003). Nodal signaling induces the midline barrier by activating Nodal expression in the lateral plate. *Development* 130, 1795–1804.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 September 2012; paper pending published: 07 October 2012; accepted: 25 October 2012; published online: 16 November 2012.

Citation: Sauer S and Klar AJS (2012) Left-right symmetry breaking in mice by left-right dynein may occur via a biased chromatid segregation mechanism, without directly involving the Nodal gene. *Front. Oncol.* 2:166. doi: 10.3389/fonc.2012.00166

This article was submitted to *Frontiers in Cancer Genetics*, a specialty of *Frontiers in Oncology*.

Copyright © 2012 Sauer and Klar. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



Discovering non-random segregation of sister chromatids: the naïve treatment of a premature discovery

Karl G. Lark*

Department of Biology, University of Utah, Salt Lake City, UT, USA

Edited by:

James L. Sherley, Boston Biomedical Research Institute, USA

Reviewed by:

Thomas Rando, Stanford University, USA

Karen Hubbard, The City College of New York, USA

Peter J. Quesenberry, Brown University, USA

*Correspondence:

Karl G. Lark, Department of Biology, University of Utah, 257 South 1400 East, Room 201, Salt Lake City, UT 84112, USA.
e-mail: lark@bioscience.utah.edu

The discovery of non-random chromosome segregation (**Figure 1**) is discussed from the perspective of what was known in 1965 and 1966. The distinction between daughter, parent, or grandparent strands of DNA was developed in a bacterial system and led to the discovery that multiple copies of DNA elements of bacteria are not distributed randomly with respect to the age of the template strand. Experiments with higher eukaryotic cells demonstrated that during mitosis Mendel's laws were violated; and the initial serendipitous choice of eukaryotic cell system led to the striking example of non-random segregation of parent and grandparent DNA template strands in primary cultures of cells derived from mouse embryos. Attempts to extrapolate these findings to established tissue culture lines demonstrated that the property could be lost. Experiments using plant root tips demonstrated that the phenomenon exists in plants and that it was, at some level, under genetic control. Despite publication in major journals and symposia (Lark et al., 1966, 1967; Lark, 1967, 1969a,b,c) the potential implications of these findings were ignored for several decades. Here we explore possible reasons for the pre-maturity (Stent, 1972) of this discovery.

Keywords: non-random segregation, sister chromatid, stem cell, mouse, radioautography

INTRODUCTION

In 1966, Richard Consigli, Harish Minocha and I published a paper in *Science* entitled "Segregation of sister chromatids in mammalian cells" (Lark et al., 1966). The first sentence of the abstract read as follows: "Segregation of sister chromatids in embryonic mouse cells in primary tissue culture is not random." In so doing, we reported the unexpected existence of non-random mitotic segregation of eukaryotic chromosomes in stem cells, the focus of this book. Non-random segregation in mouse cells was discovered as a consequence of analyzing bacterial DNA replication and segregation. During the next few years (1967–1969) we demonstrated the phenomenon in plant cells as well.

Our results are an example of the discovery of a phenomenon for which an appropriate hypothesis was lacking at the time; possibly an anomaly because it did not lend itself to any existing body of information in either a supportive or contradictory role. It was *data driven science* – and in some sense, premature (see Stent, 1972). This brief memoir traces the sequence of events leading to the discovery of non-random replication and describes the scientific context at that time: what we knew and what we did not know (or even suspect) may explain the pre-maturity that often becomes associated with data driven science.

THE DISCOVERY OF NON-RANDOM SEGREGATION

In 1963, we began a series of experiments on DNA replication (and eventually segregation) in bacteria (Lark et al., 1963; Lark and Bird, 1965; Lark, 1966b,c). A decade had passed since the announcement of the structure of DNA (Watson and Crick, 1953) during which ingenious experiments had: (i) verified that structure (Josse et al., 1961); (ii) demonstrated the existence of semi-conservative

replication in eu- and pro-karyotes (Taylor et al., 1957; Meselson and Stahl, 1958); and (iii) suggested a mechanism for regulating the initiation of DNA synthesis in bacteria (Jacob et al., 1963).

Four experimental tools were essential to these results: pulse chase as a technique for *in vivo* analysis of sequential intra cellular events (Roberts, 1964); autoradiography of tritiated thymidine labeled DNA (Taylor et al., 1957; Painter, 1958); density labeling of DNA (^{15}N : Meselson and Stahl, 1958; or 5-Bromo-uridine: Lark et al., 1963); and the use of conditional lethal mutations to dissect intracellular bacterial processes (Epstein et al., 2012).

The cell biology of bacterial growth also had been analyzed during that decade (Schaechter et al., 1958), demonstrating, among other things, that the cellular content of RNA and DNA changed when bacteria grew at different growth rates in different media. In poor media (slower growth rates), the content of RNA and DNA were lower than during more rapid growth in richer media. Our experiments utilized a strain of *Escherichia coli*, 15T⁻, which could contain either two separate replicating chromosomes per cell, or only one depending on the growth rate determined by different nutrients. When grown rapidly (in glucose), the two chromosomes would replicate at the same time, whereas at somewhat less rapid rates of growth (succinate) first one would be replicated and then the other (Lark and Lark, 1965; Lark, 1966a). As far as could be determined the two chromosomes were biologically identical, since when grown very slowly (acetate) cells contained only one chromosome, but could regenerate the two chromosome content if transferred to a medium promoting faster growth.

In order to establish the chromosome content of 15T⁻ cells grown at these different rates, we labeled cells with a pulse of tritiated thymine and then grew them in non-radioactive medium for

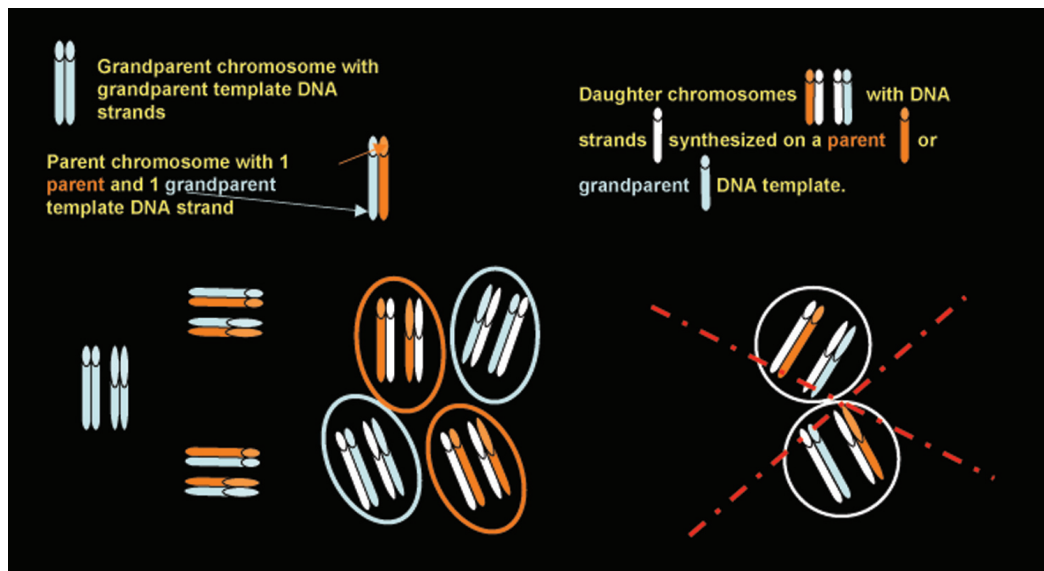


FIGURE 1 | Non-random segregation of chromosomes synthesized on DNA templates of different ages. Granddaughter cells (circled) contain chromosome sets synthesized either on grandparent DNA templates or on parent DNA template, but do not (X) receive sets that are mixtures of chromosomes synthesized some on grandparent- others on parent-templates.

different periods (chase) and plated them at different times onto non-radioactive nutrient agar to allow the development of microcolonies. Autoradiography of either cells or the microcolonies derived from these cells established the number of radioactive DNA units labeled by the pulse (Lark and Bird, 1965). We expected that after a chase period, chromosomes would be distributed randomly – i.e., these cells each containing two chromosomes would contain two radioactive chromosomes, one radioactive and one non-radioactive chromosome, or two non-radioactive chromosomes in frequencies predicted by a binomial distribution. A surprising result was that labeled and unlabeled daughter chromosomes were not distributed randomly into daughter cells. Instead, each daughter received one of each!! This suggested to us that *during DNA segregation, the cells somehow distinguished between apparently identical chromosomes on the basis of the age of their template strands* (Lark, 1966c).

We were curious if such discrimination might occur during eukaryotic somatic growth, but we had not worked with any eukaryotic system. In order to test this we decided to use a cell culture system and approached a colleague, “Dick” Consigli, who was studying polyoma virus grown on tissue cultures of mouse embryo cells (Consigli et al., 1966). In collaboration with Consigli and Minocha, we labeled *primary* cell cultures derived from embryonic tissue with tritiated thymidine and subsequently grew them in non-radioactive medium (“chase”). This serendipitous selection of a primary cell culture yielded dramatic results described in the 1966 Science paper.

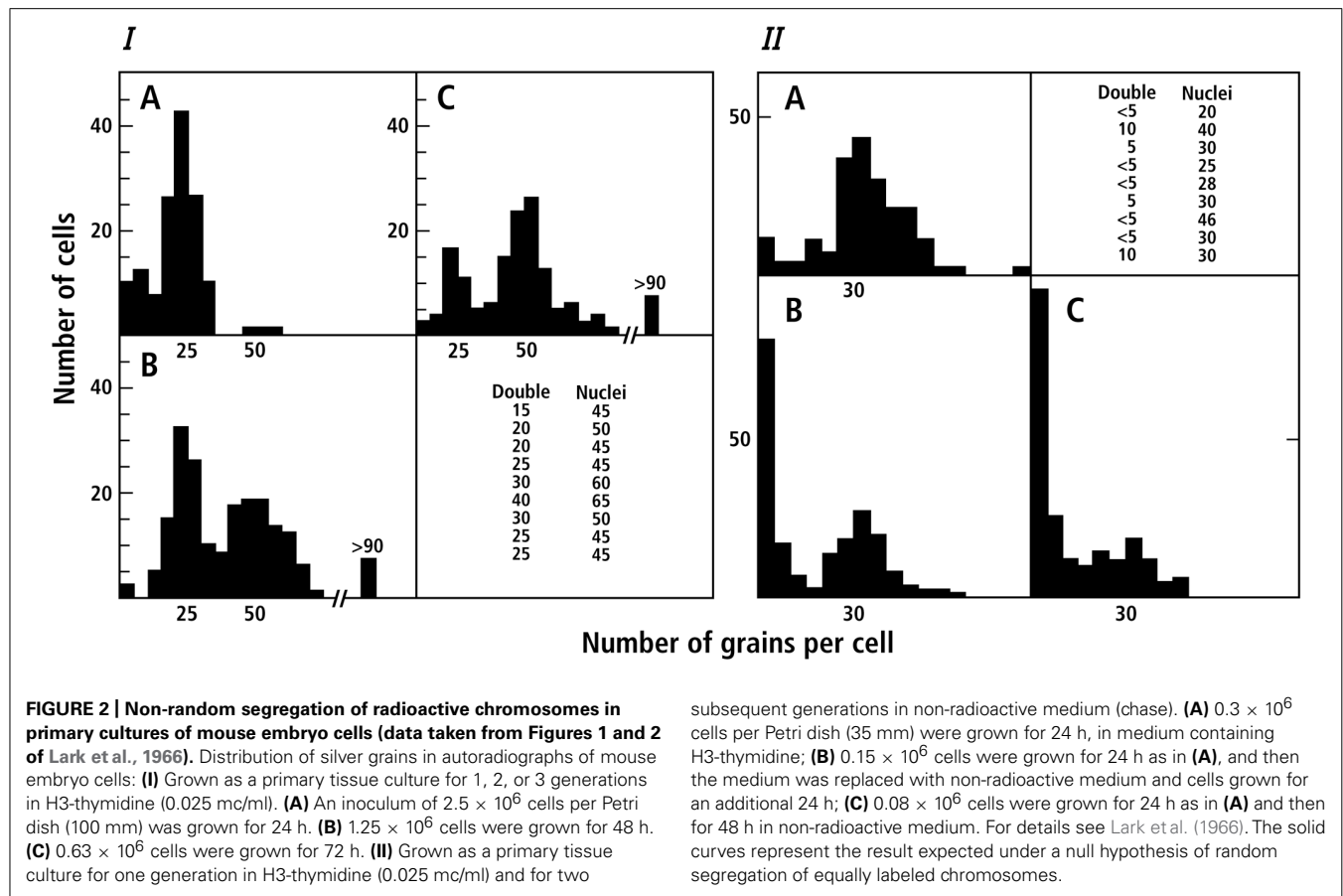
Figure 2 presents the results of labeling the cells continuously for several generations (Figure 2(I)) or of a period of radioactive labeling followed by growth in non-radioactive medium (chase Figure 2(II)). It was immediately evident that the amount of radioactivity in cells increases or decreases discontinuously in a

manner to be expected if the 40 chromosome templates that had incorporated radioactivity remained together.

An unexpected bonus of this cell system was the frequent occurrence of cells with two nuclei that had yet to divide. Radioactivity of such nuclei (grains per nucleus) are also presented in the experiments in Figures 2(I) and (II). After a two to three generation chase many of the two nucleate cells were still radioactive, but only one nucleus was heavily labeled. Segregation was not random in these mouse cells.

We soon discovered that established tissue culture lines (HeLa or CHO) had lost this property (Lark et al., 1966). Had we begun with established cell lines, we would probably have concluded that the non-random segregation we had documented in bacteria was not a property of somatic mammalian cell division, and we would have abandoned the investigation. Instead, we speculated that the polyploid nature of these established cell lines had obscured the non-random segregation of diploid chromosome sets. Although the distinction between cell lines that had acquired immortality and primary cell lines with programmed longevity was known (Hayflick and Moorhead, 1961; Hayflick, 1965), we had not considered this distinction as a possible explanation for the difference between a primary mouse line and the HeLa or CHO lines.

Our desire to avoid possible changes in ploidy during prolonged tissue culture led us to search for alternative preparations in which the non-random segregation could be studied *in vivo*. Autoradiographic analysis of mitosis in plant root tips had been used with great success by Taylor et al. (1957), who elegantly used autoradiography to establish that a chromosome was one DNA molecule that replicated semi-conservatively. We decided to use this system and settled on a pulse-chase protocol in which root tips were first grown for a period in radioactive thymidine (pulse) and then grown in non-radioactive medium for a much longer time (chase).

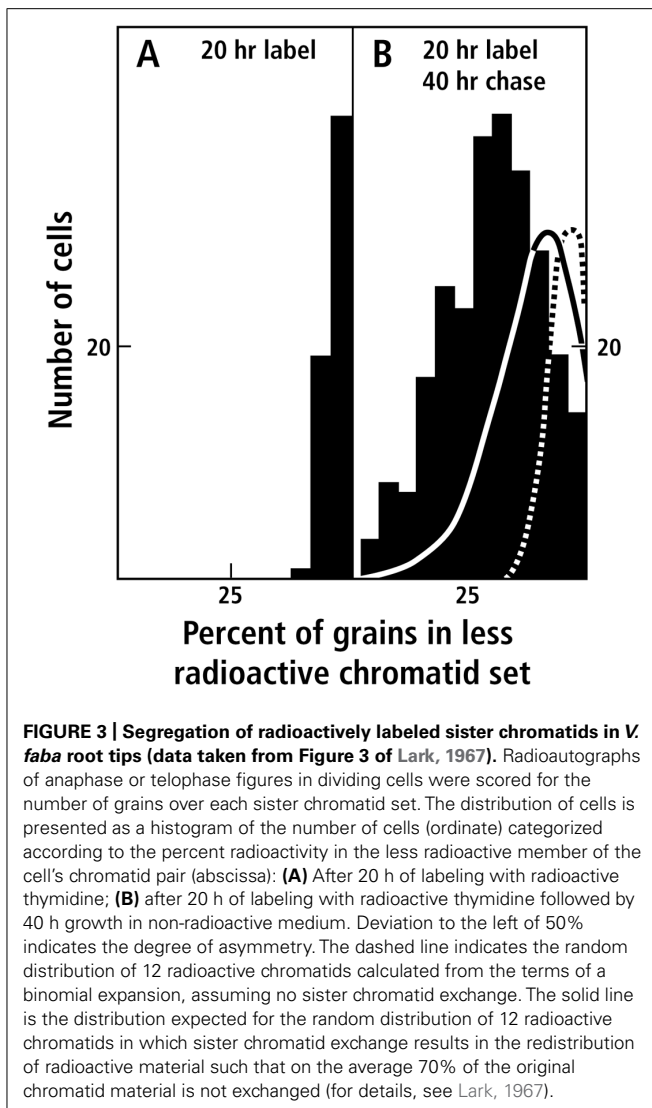


We analyzed root tips of the diploid bean, *Vicia faba* that has 12 easily visualized chromosomes, and in which sister chromatid exchange had been studied in detail by Peacock (1963). By examining anaphase preparations we could compare the amount of radioactivity in the two sets of chromosomes separating after the pulse-chase had been completed. The data (Lark, 1967), corrected for sister chromatid exchange, clearly supported non-random segregation of chromosomes in which “parent template” radioactive DNA was separated from non-radioactive “grandparent template” DNA (**Figure 3**).

To test the idea that polyploidy might obscure non-random segregation we also examined root tips of wheat, *Triticum aestivum* ($2n = 42$) a hexaploid composed of three sets of similar but not identical *homeologous* chromosomes. The results of our pulse-chase experiments were not clear-cut. The difference between chromosome sets in anaphase preparations was less apparent than in *V. faba*, (Lark, 1967), but left open the possibility that there might be a sub population of cells in which segregation was non-random. In contrast, chromosome segregation was clearly non-random in *Triticum boeoticum* ($4n = 28$; AAAA) a tetraploid relative of modern wheat. Cytological studies of mitosis in wheat (Feldman et al., 1966) had just been published suggesting that during inter-phase, chromosomes did not completely de-condense and that homeologous sets of chromosomes were not randomly distributed throughout the nucleus thus opening the possibility that segregation of each of the three diploid sets that composed

the hexaploid might be regulated autonomously. We therefore decided to further analyze segregation in *Triticum* using genetically different, but related, plants.

An outstanding achievement of 20th century evolutionary cytogenetics was the research by Hitoshi Kihara (Japan), Earnest Sears (USA), and Nikolai Vavilov (Russia; reviewed by Crow, 1994) that led to an understanding of the origins of hexaploid wheat. Their analysis established that wheat was a hexaploid composed of three diploid sets of similar, but not identical, *homeologous* chromosomes (A, B, and D; $2n = 14/\text{set}$). Their research had produced a number of genetically different *Triticum* lines. We obtained seeds from Sears and in the process learned about the history of wheat and the use of polysomic/nullisomic lines to associate phenotypes with specific chromosomes or portions of chromosomes. In polysomic/nullisomic lines of wheat, particular chromosomes, or arms of chromosomes, of one homeologous set are replaced by extra copies of the same chromosome from another, different, homeologue. For example nulli5B-tetra5D (2-5A:0-5B:4-5D) or nulli5D-tetra5B (2-5A:4-5B:0-5D) lines lack *homeologous* chromosome 5B or 5D, respectively. The results (**Figure 4**) of comparing radioactive segregation in anaphases of cells from pulse-chased root tips from these different lines led to the conclusion that a locus, or loci, on chromosome 5 regulated non-random segregation (Lark, 1969a): 5B promoted non-random segregation; 5D promoted random segregation.



Additional experiments varying the chromosome 5 dosage of different homeologs (A, B, or D) in tetraploid lines of *Triticum* (Lark, 1969a) led to the conclusions that: (i) non-random segregation was a normal process in wheat, (ii) a locus on chromosome 5D suppressed non-random segregation, (iii) the magnitude of this effect in the presence of 5B was dose dependent, and (iv) that although 5A alone (e.g., tetra 5A in *T. boeoticum*) resulted in non-random segregation, 5A was recessive to 5D in an amphidiploid (AADD = random segregation).

In summary, at the end of 1969 we knew the existence of non-random segregation in three different kingdoms: Eubacteria, Plantae, and Animalia. It appeared that in these systems, Mendel's second law was abrogated during mitosis of cells *in vivo* or in primary cell culture.

As with most data driven science, our discovery was premature in that we were unprepared for this result. In a lecture at the University of Lille in December 1854, Louis Pasteur noted that “when collecting data (*les champs de l'observation*) chance favors none but the prepared mind.” Non-random segregation

was a most unexpected finding falling on minds conditioned to the random segregation of chromosomes during meiosis. We were unprepared to consider the consequences of separating chromosome sets according to the generation of the template on which new DNA was replicated.

We did not appreciate the full implications of observing the phenomenon in bacteria and plant root tips as well as in cultured cells taken from mouse embryos and failed to explore the evolutionary implication that selection favored such a process. Had we done so, we would have concluded that: (1) maintaining an informational interactive network distributed between multiple chromosomes might be used to discriminate between cells of different generations; (2) in order for that to occur, there should be frequent changes in the chromosomes or their attached proteins, between one generation and the next (imprinting); (3) such changes could provide adaptation (for plants and bacteria) to unforeseen changes in their environment, or using programmed changes in the cell's environment (e.g., in mammals), it could facilitate morphogenesis or embryogenesis. With these last considerations in mind we might have asked, “What is special about primary mouse embryo cells?” rather than focus our attention on defects in HeLa and CHO cells.

Significantly two events had occurred in the late 1960s that should have directed my attention to the important role that non-random segregation might play in biological systems: One was the demonstration by my wife Cynthia Lark (Lark, 1968a,b,c) of the methylation of newly synthesized strands of bacterial DNA and the discovery that DNA replication would not proceed *in vivo* if a template strand was not methylated. Most importantly, she collaborated with Werner Arber to demonstrate that host specificity phenotypes were regulated by *in vivo* methylation of DNA (Lark and Arber, 1970) – i.e., *modification of DNA without changing nucleotide sequence could result in changes of phenotype, an early manifestation of imprinting*.

The other was a visit with Don Brown at the Carnegie Institute of Embryology in Baltimore. When I told Brown about non-random segregation in mouse cells he responded that if this were true it would have extremely important ramifications for developmental biology.

As a cell biologist I had naively concentrated on a cellular and genetic description of the phenomenon, but failed to even speculate on the possible advantages conferred on an organism by this aspect of mitosis. Had I done so, I would have realized that this discovery signaled changes involving chromosomes or closely associated (tightly bound) proteins that differentiated one generation of sister chromatids from another, a realization that should have triggered an attempt to identify such *epigenetic* changes.

RETROSPECTIVE

This research was interrupted in 1969 by illness and a change in my career to become chair of the Department of Biology at the University of Utah. Research on bacteria continued in my laboratory carried out by post-docs and students, but personal, hands on, experimentation ceased for a period of several years.

The department in Utah was, and still is, united by a common deep interest in evolutionary phenomena and my exposure to this from different faculty, ranging from population ecologists

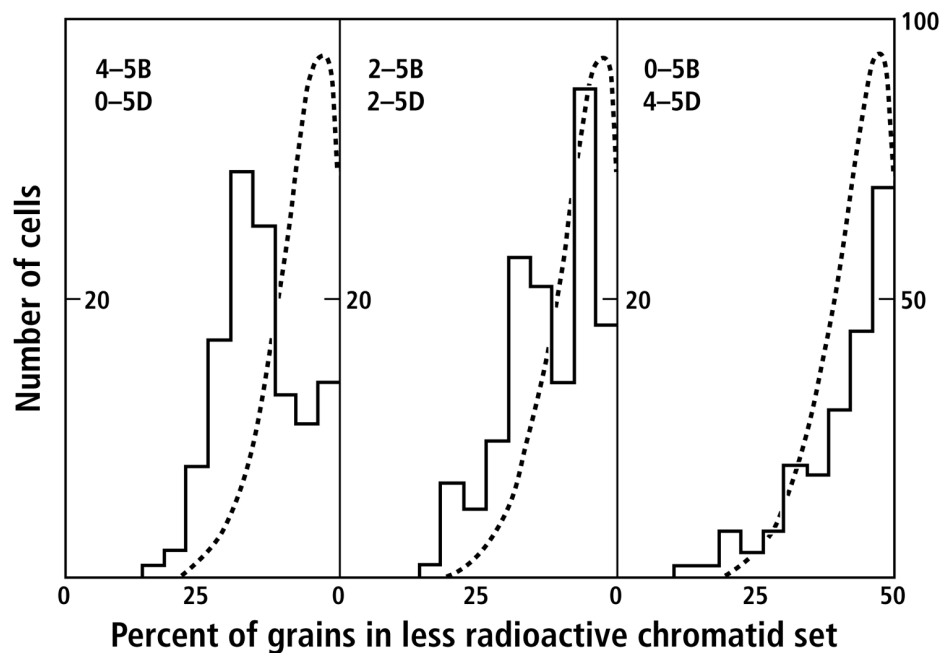


FIGURE 4 | Sister chromatid segregation in mitotic figures of *T. aestivum*: effects of varying the dosage of chromosomes 5B and 5D (data from Lark, 1969a, Figure 7). Histograms of the frequency of cells in which the indicated percentages of grains were observed in the less

radioactive chromatid set during anaphase or telophase. Deviation to the left of 50% indicates the degree of asymmetry. The dashed curve represents the expected binomial distribution for random segregation (for details, see Lark, 1969a).

to molecular biologists, returned my thoughts if not my hands to non-random segregation. John Cairns, a friend and colleague, who shared a deep interest in DNA replication, visited for a few weeks in 1972 and 1973, during which we discussed the non-random segregation experiments. Our results interested him and eventually led to his hypothesis on the role of non-random segregation in maintaining template fidelity, consequently reducing the frequency of cancer in populations of rapidly reproducing intestinal epithelial cells (Cairns, 1975). Cairns' ideas thus were the first formal exposition of the concept that in higher eukaryotes non-random segregation would serve to differentiate between cells of different generations whose function and fate would thenceforth be different.

At that time very little genetic analysis of interactive networks had been carried out. Genetics was still focused on single gene effects. Phenotypes were mostly qualitative and the techniques for analyzing quantitative genetic systems were just beginning to be developed (e.g., Falconer and Mackay, 1996). Today, we are

very aware of genome wide interactions and their importance in regulating a multitude of quantitative phenotypes. Thus, the selective advantage of non-random mitotic chromosome segregation in preserving networks of inter-chromosome information now seems evident. Cairns' ideas focused on the exclusion of deleterious mutations thus maintaining the integrity of stem cells with "healthy" genomes. However, if we invoke the idea of directed epigenetic variation involving multiple chromosomes, non-random segregation can be viewed as a mechanism for preserving cell lineages with useful adaptations. Such lineages could then be maintained throughout development and morphogenesis (Reik, 2007; Fedoriw et al., 2012). The ability to maintain such networks also would provide flexibility in coping with environmental variation constituting a powerful selective advantage for plants as well as for animals that lack intrinsic homeostatic control (e.g., see Feil and Berger, 2007). These, as well as other consequences, promise an exciting future for the investigation of phenotypes benefiting from non-random chromosome segregation during mitosis.

REFERENCES

- Cairns, J. (1975). Mutation selection and the natural history of cancer. *Nature* 255, 197–200.
- Consigli, R. A., Minocha, H. C., and Abo-Ahmed, H. C. (1966). Multiplication of polyoma virus. II. Source of constituents for viral deoxyribonucleic acid and protein synthesis. *J. Bacteriol.* 92, 789–791.
- Crow, J. F. (1994). Hitoshi Kihara, Japan's pioneer geneticist. *Genetics* 137, 891–894.
- Epstein, R. H., Bolle, A., and Steinberg, C. M. (2012). Amber mutants of bacteriophage T4D: their isolation and genetic characterization. *Genetics* 190, 831–840.
- Falconer, D. S., and Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics*, Ed 4. Longmans Green, Harlow, Essex, UK.
- Fedoriw, A., Mugford, J., and Magnusson, T. (2012). Genomic imprinting and epigenetic control of development. *Cold Spring Harb. Perspect. Biol.* 4, a008136.
- Feil, R., and Berger, F. (2007). Convergent evolution of genomic imprinting in plants and mammals. *Trends Genet.* 23, 192–199.
- Feldman, M., Mello-Sampayo, T., and Sears, E. R. (1966). Somatic association in *Triticum aestivum*. *Proc. Natl. Acad. Sci. U.S.A.* 56, 1192–1199.
- Hayflick, L. (1965). The limited in vitro lifetime of human diploid cell strains. *Exp. Cell Res.* 37, 614–636.
- Hayflick, L., and Moorhead, P. S. (1961). The serial cultivation of human diploid cell strains. *Exp. Cell Res.* 25, 585–621.
- Jacob, F., Brenner, S., and Cuzin, F. (1963). On the regulation of DNA replication in bacteria. *Cold*

- Spring Harb. Symp. Quant. Biol.* 28, 329–348.
- Josse, J., Kaiser, A. D., and Kornberg, A. (1961). Enzymatic synthesis of deoxyribonucleic acid: VIII. Frequencies of nearest neighbor base sequences in deoxyribonucleic acid. *J. Biol. Chem.* 236, 864–875.
- Lark, C. (1966a). Regulation of deoxyribonucleic acid synthesis in *Escherichia coli*: dependence on growth rates. *Biochim. Biophys. Acta* 119, 517–525.
- Lark, K. G. (1966b). “Chromosome replication in *Escherichia coli*,” in *Cell Synchrony*, eds. I. L. Cameron, and G. M. Padilla (New York: Academic Press), 54–80.
- Lark, K. G. (1966c). Regulation of chromosome replication and segregation in bacteria. *Bacteriol. Rev.* 32, 3–32.
- Lark, C. (1968a). Studies on the in vivo methylation of DNA in *Escherichia coli* 15T[−]. *J. Mol. Biol.* 31, 389–399.
- Lark, C. (1968b). Effect of the methionine analogs, ethionine and norleucine, on DNA synthesis in *Escherichia coli* 15T[−]. *J. Mol. Biol.* 31, 401–414.
- Lark, C. (1968c). Methylation of DNA and the replication of the bacterial chromosomes, in *Some Recent Developments in Biochemistry*, eds. H.-W. Li, P. O. P. Ts'o, P. C. Huang, and T. T. Kuo (Peking: Academia Sinica), 104–111.
- Lark, C., and Arber, W. (1970). Host specificity of DNA produced by *Escherichia coli*. 13. Breakdown of cellular DNA upon growth in ethionine of strains with r plus-15, r plus-P1 or r plus-N3 restriction phenotypes. *J. Mol. Biol.* 52, 337–348.
- Lark, K. G. (1967). Nonrandom segregation of sister chromatids in *Vicia faba* and *Triticum boeoticum*. *Proc. Natl. Acad. Sci. U.S.A.* 58, 352–359.
- Lark, K. G. (1969a). Sister chromatid segregation during mitosis in polyploidy. *Wheat Genetics* 62, 289–305.
- Lark, K. G. (1969b). “DNA replication and the segregation of chromosomes,” in *Some Recent Developments in Biochemistry*, eds. H.-W. Li, P. O. P. Ts'o, P. C. Huang, and T. T. Kuo (Peking: Academia Sinica), 55–63.
- Lark, K. G. (1969c). “Non-random segregation of sister chromatids,” in *Genetics and developmental biology*, ed. H. J. Teas (Lexington: University of Kentucky Press) 8–24.
- Lark, K. G., and Bird, R. E. (1965). Segregation of the conserved units of SNA in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 54, 1444–1450.
- Lark, K. G., Consigli, R. A., and Minocha, H. C. (1966). Segregation of sister chromatids in mammalian cells. *Science* 154, 1202–1205.
- Lark, K. G., Eberle, H., Consigli, R. A., Minocha, H. C., Chai, N., and Lark, C. (1967). “Chromosome segregation and the regulation of DNA replication,” in *Organizational Biosynthesis*, eds. H. J. Vogel, J. O. Lampen, and V. Bryson (New York: Academic Press), 63–89.
- Lark, K. G., and Lark, C. (1965). Regulation of chromosome replication in *Escherichia coli*: alternate replication of two chromosomes at slow growth rates. *J. Mol. Biol.* 13, 105–126.
- Lark, K. G., Repko, T., and Hoffman, E. J. (1963). The effect of amino acid deprivation on subsequent deoxyribonucleic acid replication. *Biochim. Biophys. Acta* 76, 9–24.
- Meselson, M., and Stahl, F. W. (1958). The replication of DNA in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 44, 671–682.
- Painter, R. B. (1958). Cellular proliferation in the mouse as revealed by autoradiography with tritiated thymidine. *Proc. Natl. Acad. Sci. U.S.A.* 44, 476–483.
- Peacock, J. (1963). Chromosome duplication and structure as determined by autoradiography. *Proc. Natl. Acad. Sci. U.S.A.* 49, 793–801.
- Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447, 425–432.
- Roberts, R. B. (1964). *Studies of Macromolecular Biosynthesis*. Washington, DC: Carnegie Institution of Washington, publication 624.
- Schaechter, M., MaalØe, O., and Kjeldgaard, N. O. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *J. Gen. Microbiol.* 19, 592–606.
- Stent, G. S. (1972). Prematurity and uniqueness in scientific discovery. *Sci. Am.* 227, 84–93.
- Taylor, J. H., Woods, P. S., and Hughes, W. L. (1957). The organization and duplication of chromosomes as revealed by autoradiographic studies using tritium-labeled thymidine. *Proc. Natl. Acad. Sci. U.S.A.* 43, 122–128.
- Watson, J. D., and Crick, F. H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171, 737–738.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 November 2012; accepted: 19 December 2012; published online: 01 February 2013.

Citation: Lark KG (2013) Discovering non-random segregation of sister chromatids: the naïve treatment of a premature discovery. *Front. Oncol.* 2:211. doi: 10.3389/fonc.2012.00211

This article was submitted to *Frontiers in Cancer Genetics*, a specialty of *Frontiers in Oncology*.

Copyright © 2013 Lark. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.



Comparison of the transcriptomes of long-term label retaining-cells and control cells microdissected from mammary epithelium: an initial study to characterize potential stem/progenitor cells

Ratan K. Choudhary¹, Robert W. Li², Christina M. Evock-Clover² and Anthony V. Capuco^{1,2,*†}

¹ Department of Animal and Avian Sciences, University of Maryland, College Park, MD, USA

² Bovine Functional Genomics Laboratory, USDA-ARS, Beltsville, MD, USA

Edited by:

James L. Sherley, Boston Biomedical Research Institute, USA

Reviewed by:

James L. Sherley, Boston Biomedical Research Institute, USA

Minsoo Noh, Ajou University, South Korea

*Correspondence:

Anthony V. Capuco, Bovine Functional Genomics Laboratory, Building 200, Room 14, BARC-East, Beltsville, MD 20705, USA.

e-mail: tony.capuco@ars.usda.gov

[†] Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. The USDA is an equal opportunity provider and employer.

Background: Previous molecular characterizations of mammary stem cells (MaSC) have utilized fluorescence-activated cell sorting or *in vitro* cultivation of cells from enzymatically dissociated tissue to enrich for MaSC. These approaches result in the loss of all histological information pertaining to the *in vivo* locale of MaSC and progenitor cells. Instead, we used laser microdissection to excise putative progenitor cells and control cells from their *in situ* locations in cryosections and characterized the molecular properties of these cells. MaSC/progenitor cells were identified based on their ability to retain bromodeoxyuridine for an extended period.

Results: We isolated four categories of cells from mammary epithelium of female calves: bromodeoxyuridine label retaining epithelial cells (LREC) from basal (LRECb) and embedded layers (LRECe), and epithelial control cells from basal and embedded layers. Enriched expression of genes in LRECb was associated with stem cell attributes and identified WNT, TGF- β , and MAPK pathways of self renewal and proliferation. Genes expressed in LRECe revealed retention of some stem-like properties along with up-regulation of differentiation factors.

Conclusion: Our data suggest that LREC in the basal epithelial layer are enriched for MaSC, as these cells showed increased expression of genes that reflect stem cell attributes; whereas LREC in suprabasal epithelial layers are enriched for more committed progenitor cells, expressing some genes that are associated with stem cell attributes along with those indicative of cell differentiation. Our results support the use of DNA label retention to identify MaSC and also provide a molecular profile and novel candidate markers for these cells. Insights into the biology of stem cells will be gained by confirmation and characterization of candidate MaSC markers identified in this study.

Keywords: label retention, mammary stem cells, mammary progenitor cells, stem cell markers, laser microdissection

INTRODUCTION

In female mammals, growth and development of mammary glands occur primarily postnatally, with mammary function in the mature animal being tightly coupled to reproductive strategy. This dictates cycles of mammary growth, differentiation, lactation, and regression, during which mammary stem cells (MaSC) provide for the lineages of luminal and basal (myoepithelial) epithelial cells in the ducts and alveoli. Although mice have provided the primary model for study of mammary growth and development, a single model species cannot provide comprehensive knowledge. Because mammary glands of prepubertal calves have a tissue architecture resembling that of the prepubertal human breast more closely than does mouse (Capuco et al., 2002), cows provide an additional experimental model for human breast development. Increased knowledge of MaSC is directly applicable to agriculture

and the development of management schemes to enhance the lifetime productivity of dairy cows and other species.

A method that has been used to identify MaSC is based upon the capacity of these cells to retain 5-bromo-2'-deoxyuridine (BrdU) labeled DNA for an extended period (Kenney et al., 2001; Welm et al., 2002; Smith, 2005; Capuco, 2007). Retention of labeled DNA strands may be attributed to the ability of stem cells to retain the parental DNA strand during asymmetric cell division (Cairns, 1975) or to quiescence of the stem cell population such that the DNA label is not diluted by frequent cell divisions (Klein and Simons, 2011). During rapid mammary growth in the mouse, label retaining epithelial cells (LREC) appear to retain label by asymmetric distribution of DNA strands, as evidenced by a rapid proliferation index of the LREC (Smith, 2005). During periods of low mammary proliferation, quiescence of the stem cell

population may account for retention of label. LREC are enriched in populations that exhibit MaSC capacity, i.e., the ability to regenerate mammary epithelium upon transplantation into the cleared mammary fat pad of syngeneic mice (Welm et al., 2002).

We previously reported that LREC in mammary epithelium of calves were localized in the basal layer (LRECb) and in the embedded (LRECe) layers between the basal and luminal cells of a multi-layered epithelium (Capuco, 2007; Capuco et al., 2009). The LREC in bovine mammary gland appeared to have a modest proliferation rate in which 5.4% of LREC co-expressed Ki-67 (Capuco, 2007). LRECb were estrogen receptor- α (ESR1) -negative and hypothesized to be MaSC, whereas the LRECe were a mixed population of ESR1-positive and -negative cells that were hypothesized to be progenitor cells (Capuco, 2007; Capuco et al., 2009). The estrogen receptor status of MaSC is of considerable interest because of the importance of estrogens for MaSC function, mammary ductal growth, and tumorigenesis. MaSC of mouse and human are ESR1-negative (Anderson and Clarke, 2004; Asselin-Labat et al., 2006; Sleeman et al., 2007; Lamarca and Rosen, 2008).

Morphological evidence suggests that MaSC are basally localized within the mammary epithelium, typically underlain by cytoplasmic extensions of epithelial cells and in close proximity to ESR1-positive epithelial cells (Smith and Chepko, 2001; Briskin and Duss, 2007). However, MaSC have not been fully characterized due to technical limitations inherent in stem cell identification and in isolation of cells from known locations within the mammary epithelium. Based on fluorescence-activated cell sorting with multiple biomarkers and use of mammary transplantation methods to evaluate multi-lineage potency, Shackleton, Stingl, and colleagues obtained and characterized a population of cells, from enzymatically dispersed mammary tissue, that was enriched for MaSC (Shackleton et al., 2006; Stingl et al., 2006). Critical to the success of this pioneering approach was use of markers to deplete the population of hematopoietic (CD45 and TER119) and endothelial cells (CD31), as well as markers to select epithelial cells (CD29, CD49f), likely from a basal location, that expressed heat stable antigen (CD24). Another approach utilized for enrichment and characterization of human MaSC involved characterization of mammary epithelial cells that possess multipotency potential *in vitro* (Dontu et al., 2003).

Cell sorting techniques have also been applied to suspensions of bovine mammary cells in an attempt to enrich for MaSC. Motyl et al. (2011) isolated and evaluated gene expression in a population of mammary cells that were isolated on the basis of SCA1 expression and showed up-regulation of genes that are characteristic of hematopoietic cells. However, because accompanying micrographs clearly show that most SCA1-positive cells were in the mammary stroma and methods to enrich for mammary epithelial cells were not employed, the gene expression profile likely cannot be attributed to MaSC. Furthermore, previous research indicates the likelihood of hematopoietic cells populating the mammary stem cell niche is highly unlikely (Niku et al., 2004). Research by Martignani et al. (2010) utilized aldehyde dehydrogenase (ALDH) activity as a selection criterion for cell sorting and demonstrated that cells with low ALDH activity were capable of regenerating functional structures of mammary epithelium within collagen gels implanted beneath the kidney capsule of immunodeficient mice.

This latter study not only provides data pertaining to characteristics of bovine bipotent progenitor cells, but validates a means to assess such potency. Most recently, Rauner and Barash (2012) used the multiparameter cell sorting technique developed for enrichment of murine MaSC (Shackleton et al., 2006) to obtain and characterize four populations of mammary epithelial cells from dissociated bovine mammary gland. The differentiation and growth potential of the cells were assessed by *in vitro* colony formation and mammosphere assays. This study confirmed many of the general aspects of MaSC/progenitor cells evident in mouse and human studies. The four populations included putative bovine MaSC (CD24^{med}CD49f^{pos}) that were bipotent (myoepithelial and luminal) and possessed a high growth rate; basal bipotent progenitors with medium growth rate and low sphere generating potential; luminal unipotent progenitors with low growth rate; and luminal unipotent cells with very limited proliferative activity. Although putative MaSC typically possessed little or no ALDH activity, as reported previously (Martignani et al., 2010), 0.4% of total viable cells expressed high ALDH activity, which they hypothesized represent the MaSC population.

In addition to issues pertaining to the isolation of MaSC from a mixed suspension of mammary cells, all previous studies have evaluated MaSC after removing them from their stem cell niche, i.e., the microenvironment of surrounding signaling molecules and other non-cellular components that support stem cell function and survival. We have taken an approach that retains histological information by characterizing gene expression in putative MaSC directly after their *in situ* excision from the mammary epithelium. The histological location of all cells interrogated was known.

In the present study, putative stem and progenitor cells (LREC) were identified and excised from cryosections using laser microdissection. It must be recognized that identification of putative MaSC and progenitor cells on the basis of long-term retention of DNA label is to select the cells based upon their life-history (i.e., the extent of label retention represents an integration of the cell's past proliferation and differentiation events). Consequently, one would anticipate that selecting putative MaSC and progenitor cells based on label retention is likely to represent enrichment for these cell populations. In this study, LREC and neighboring epithelial control (non-LREC) cells were excised from two different locations: basal and embedded layers of the mammary epithelium. We hypothesized that LRECb are enriched for MaSC whereas LRECe are enriched for more committed progenitor cells, and that by comparing the transcriptomes of these cells with neighboring control cells we would obtain molecular profiles and biomarkers for MaSC and progenitor cells. Results are consistent with these hypotheses and provide novel candidate markers for MaSC and progenitor cells.

MATERIALS AND METHODS

EXPERIMENTAL ANIMALS AND MAMMARY TISSUE

Use of animals for this study was approved by the Beltsville Agricultural Research Center's Animal Care and Use Committee. Tissues for this study were obtained from five Holstein heifers at approximately 5 months of age (4.8 ± 0.05 , mean \pm SE). At approximately 3 months of age, heifers were injected intravenously with BrdU (Sigma-Aldrich Co., St. Louis, MO, USA) for five

consecutive days. BrdU was administered in a saline solution containing 20 mg BrdU/ml (0.9% sodium chloride; pH 8.2) at a dosage of 5 mg/kg body weight, as described previously (Capuco, 2007). Heifers were sacrificed humanely at the Beltsville Agricultural Research Center abattoir 45 days after the last BrdU injection. Mammary tissue (~5 mm × 5 mm × 5 mm) was collected from the outer parenchymal region (region in close proximity to the border with mammary fat pad) of a rear mammary gland. Individual samples were immediately embedded in OCT compound (Sakura, Torrance, CA, USA), frozen in liquid nitrogen vapor and stored at -80°C until use.

Cryosections of 8 μm thickness were thaw-mounted on ultraviolet-irradiated PEN slides (Leica AS, Wetzlar, Germany) and stored at -80°C until BrdU immunostaining and laser microdissection within 8 days. Mammary tissues harvested for histological validation of microarray data were fixed overnight in 10% neutral buffered formalin at 4°C and then stored in 70% ethanol until further processing. Tissues were then dehydrated and embedded in paraffin according to standard techniques and sectioned at 5 μm thickness onto Superfrost-plusTM slides (Erie Scientific Co., Portsmouth, NH, USA).

BrdU IMMUNOSTAINING TO IDENTIFY PUTATIVE MaSC

Putative MaSC were identified as those cells in cryosections that retained BrdU label (**Figure 1D**), visualized using an optimized method for BrdU immunostaining that retains RNA quality in tissue cryosections (Choudhary et al., 2010). Sections were individually processed immediately before laser microdissection. The cryosections were fixed in acetone/polyethylene glycol 300 (9:1 v/v) at -20°C for 2 min and air dried for 1 min and then incubated with 0.5% methyl green for 2 min at room temperature (RT). After a brief wash (10 s) with nuclease-free phosphate buffered saline (nfpBS), 400 μl of a pre-warmed solution of 70% deionized formamide in nfpBS was pipetted onto the tissue and the section incubated at 60°C for 4 min. The section was washed with antibody dilution buffer (nfpBS with 1% normal goat serum and 0.1% triton-X 100) at 4°C on a metal plate kept on ice to prevent re-annealing of DNA strands and then incubated with mouse monoclonal anti-BrdU antibody conjugated to Alexa 488 (Clone PRB-1, 1:10 dilution, Molecular Probes, Carlsbad, CA, USA) for 5 min at RT in the dark. The section was washed briefly before counterstaining with propidium iodide (2.5 $\mu\text{g}/\mu\text{l}$ in nfpBS). Finally, the slide was washed with nuclease-free water (10 s), dehydrated in ascending concentrations of ethanol and air dried before laser microdissection.

LASER MICRODISSECTION AND cDNA AMPLIFICATION

Immediately after staining, sections were examined and cells excised with a laser microdissection system equipped for epifluorescence microscopy (Leica AS-LMD, Mannheim, Germany). The laser setting was determined empirically and dissection performed using the 40 \times objective. We dissected 6–13 cells per category per heifer. For each animal, cells in a given category were collected into the cap of a 0.2 ml thin-walled PCR tube (Biozyme Scientific GmbH, Hess Oldendorf, Germany). Total processing time for immunostaining and microdissection was less than 1 h, and only one slide was processed at a time. Four categories of cells

were dissected: LREC from basal (LRECb) and embedded layers (LRECe), and epithelial control cells from basal (ECb) and embedded layers (ECe). Cells within the cap were dissolved in 2 μl of lysis buffer (WT-OvationTM One-Direct RNA Amplification System; NuGEN Technologies, Inc., San Carlos, CA, USA). The tube was capped and centrifuged for 1 min at $14,000 \times g$, after which the tube and contents were vortexed gently for 30 s and centrifuged briefly before placing on ice. First stand cDNA synthesis and amplification reaction were carried out using Ribo-SPIA-based methodology according to the manufacturer's recommendations. Concentrations of amplified cDNA were determined spectrophotometrically (ND-1000, NanoDrop Technologies, Rockland, DE, USA). A known amount of high quality RNA (250 pg) was used as positive control for cDNA amplification. Nuclease-free water was used as a no-template control for cDNA amplification. The amplified cDNA was evaluated using RNA Nano-chips to estimate the median fragment size (Agilent Technologies, Palo Alto, CA, USA). Median fragment size for amplified samples was similar to the positive control and fell within the expected range of 100–300 bp, whereas products for the no-template control were <50 bp.

MICROARRAY ANALYSIS

Oligonucleotide microarray analysis was performed using a custom bovine microarray (Nimblegen, Inc., Madison, WI, USA) as described previously (Li et al., 2006). The bovine microarray consisted of 86,191 unique 60-mer oligonucleotides, representing 45,383 bovine sequences. The array design was based upon a TIGR assembly (release 11.0 from 2004). However, all 60-mer oligonucleotides on the array were annotated against current bovine RefSeq databases as well as the latest version of ENSEMBL bovine gene build v65.0 (released on December 2011¹). After hybridization, scanning, and image acquisition, the data were extracted from the raw images using NimbleScan software (NimbleGen). A total of 21 microarrays (five animals \times four categories of cells, and no-template amplification control) were used. Relative signal intensities (\log_2) for each feature were generated using the robust multi-array average algorithm (Irizarry et al., 2003) and data were processed based on the quantile normalization method (Bolstad et al., 2003). Only oligos that provided hybridization signal intensities for samples that exceeded $3 \times$ the signal intensity obtained with the no-template amplification control (water blank) were included in the analysis. Furthermore, only sample signal intensities exceeding twice the array background intensity (mean of lowest 3% of oligo intensities) were considered for analysis.

P values were calculated using a modified *t*-test. Fold changes were calculated as the ratio of the means of background-adjusted, normalized fluorescent intensity of cells of interest to their respective controls. Group-wise comparisons were performed in accordance with recommendations of the Microarray Quality Control project (Shi et al., 2006, 2008) based on *t*-test ($P < 0.05$) followed by fold change (twofold as a cutoff) to determine significance. These criteria were shown to achieve a balance of reproducibility, sensitivity, and specificity, using single or multiple microarray platforms (Shi et al., 2006, 2008; Chin et al., 2009, 2010; Wang

¹<http://uswest.ensembl.org/downloads.html>

et al., 2011). Based upon these criteria, genes that were differentially expressed were then subjected to pathway analysis (IPA, Ingenuity Systems²).

The microarray data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and are accessible through GEO Series accession number GSE31541³.

REALTIME QUANTITATIVE RT-PCR

Realtime quantitative RT-PCR (qRT-PCR) was performed using aliquots of amplified cDNA from all animals and an IQ SYBR Green Supermix kit (Bio-Rad Laboratories, Hercules, CA, USA). Each reaction was performed in a 25 μ l reaction volume containing 200 nM of each amplification primer and 2 ng of cDNA. The amplification was performed in a Bio-Rad iCycler using the following protocol: 95°C – 60 s; 45 cycles of 94°C – 15 s, 61°C – 30 s, and 72°C – 30 s. A melting curve analysis was performed for each primer pair. Standards were prepared from PCR amplicons purified using the QIAquick purification kit (Qiagen Inc., Valencia, CA, USA). Product concentrations were determined using the Agilent 2100 BioAnalyzer and DNA 500 kits (Agilent Technologies) and diluted to contain 1×10^2 to 1×10^8 molecules/ μ l. Quantity of cDNA in unknown samples was calculated from the appropriate external standard curve run simultaneously with samples.

IMMUNOHISTOCHEMISTRY

Paraffin sections were dewaxed in xylene and hydrated in a graded series of ethanol to phosphate buffered saline (PBS, pH 7.4). Tissue sections were quenched with 3% H₂O₂ in PBS for 10 min and then washed in PBS. Antigen retrieval was performed by incubation with 70% formamide in PBS at 60°C for 5 min, or microwave heating in 10 mM Tris containing 1 mM EDTA, pH 9.0 (5 min heat, 5 min rest, 5 min heat, 25 min cooling). Sections were blocked with casein (CAS-block™, Invitrogen, Carlsbad, CA, USA). Primary antibodies NR5A2, NUP153, and HNF4A (Abcam Inc., Cambridge, MA, USA) were used at 1:200 dilution and FNDC3B (Santa Cruz, Santa Cruz, CA, USA) at 1:50. Sections were incubated with primary antibody for 2 h at RT or overnight at 4°C. After washing in PBS, sections were incubated with horseradish peroxidase-conjugated broad spectrum secondary antibody (ImmPRESS anti-mouse/anti-rabbit, Vector Labs, Burlingame, CA, USA). Positively labeled cells were visualized brown or purple using 3,3'-diaminobenzidine or ImmPACT VIP (Vector Labs), respectively. Slides were washed and then counterstained with hematoxylin or methyl green.

To determine if cells expressing FNDC3B were LREC, dual antigen labeling was performed. Tissue sections were processed as described earlier and incubated with mouse monoclonal BrdU antibody (Clone BMC 9318, 2 μ g/ml; Roche Diagnostics Corp., Indianapolis, IN, USA) for 2 h at RT. Sections were then incubated with vector ImmPRESS anti-mouse polymer detection reagent (Vector Labs) for 20 min, followed by washing in PBS. BrdU was detected by incubation for 10 min with the chromagen 3,3'-diaminobenzidine. Sections were then washed in deionized water.

Peroxidase activity was quenched for a second time with 3% H₂O₂ in PBS, followed by washings with water. Sections were blocked with casein and then incubated overnight at 4°C with FNDC3B rabbit polyclonal antibody (1:50 dilution), washed, and then incubated with horseradish peroxidase-conjugated broad spectrum secondary antibody. Sections were washed with PBS and FNDC3B staining was visualized after incubation with a contrast purple chromogen, ImmPACT™ VIP peroxidase substrate (Vector Labs). Sections were washed in deionized water, counterstained with 0.5% aqueous methyl green (Vector Labs), differentiated in 0.05% acetic acid/acetone, washed dehydrated in ethanol, cleared in xylene, and mounted in DPX (Sigma). Omission of primary antibodies was used for negative controls.

Immunofluorescence staining was performed, as described previously (Capuco, 2007), to determine the ESR1 status of LREC by assessing the co-localization of BrdU and ESR1.

RESULTS

IDENTIFICATION OF LREC IN THE TERMINAL DUCTULAR UNITS OF BOVINE MAMMARY GLAND

During the period of ductal morphogenesis, the prepubertal mammary gland grows allometrically and mammary ducts expand into the surrounding mammary fat pad (Capuco et al., 2002; Meyer et al., 2006). The terminal ductular units of the prepubertal mammary gland, which are prevalent at this time, are arborescent structures composed of a multilayered epithelium (Capuco et al., 2002; **Figures 1A,B**). One approach, which we have utilized, to identify putative stem cells is based on the observation that somatic stem cells often retain labeled DNA strands for a prolonged period after initial labeling with tritiated thymidine or BrdU (Potten et al., 1978; Bickenbach, 1981). In mice, intestinal crypt cells (Potten et al., 2002), muscle satellite cells (Conboy et al., 2007), and putative MaSC (Welm et al., 2002; Smith, 2005) retain labeled DNA. Although long-term retention of BrdU does not appear to be a universal marker for somatic stem cells, it appears to provide a means for identifying putative stem/progenitor cells in mammary gland. After staining BrdU-labeled cells in cryosections without compromising RNA quality (Choudhary et al., 2010), we employed laser microdissection to collect LREC from basal and embedded layers of the mammary epithelium, along with appropriate control cells (**Figures 1C,D**). The transcriptome of these cells was interrogated by microarray analysis, from which we based our characterization of these interesting LREC in bovine mammary gland.

TRANSCRIPTOMES OF LRECb vs. ECb

To evaluate the hypothetical stem cell nature of LRECb, we compared the transcript profiles of LRECb vs. neighboring control cells (ECb). This analysis identified 605 genes that were differentially expressed between these two cell types (Table S1 in Supplementary Material). Of these, 476 corresponded to genes that were functionally annotated in the Ingenuity Pathway Analysis database. Differentially expressed genes were involved in pathways linked to cancer, gene expression, cell growth and proliferation, and cell death (Table S2 in Supplementary Material). A number of genes with documented relevance to MaSC were identified in this analysis (**Tables 1 and 2**). Low expression of *ESR1* and

² www.ingenuity.com

³ http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE31541.

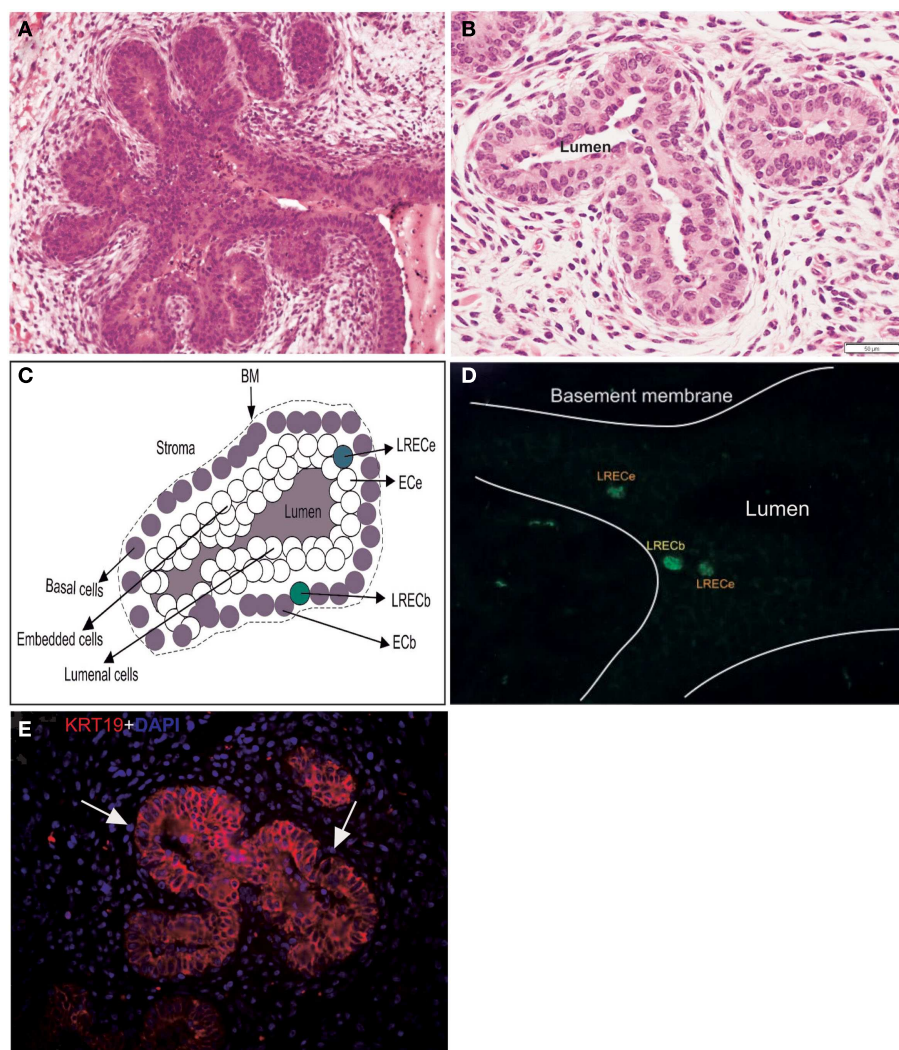


FIGURE 1 | Characteristics and dissection of the mammary epithelium of a prepubertal heifer. (A,B) Micrographs of the terminal ductal units of a prepubertal bovine mammary gland illustrating the arborescent nature of these terminal ducts and their multilayered epithelium. **(C)** Diagrammatic representation of the mammary epithelium illustrating the four classes of

cells dissected. BM, basement membrane. **(D)** Micrograph of cryosections stained for BrdU antigen prior to laser microdissection, localization of LRECb and LRECE are depicted. **(E)** Micrograph of epithelium of terminal ductal unit stained for KRT19. The basal layer consists of cells that are predominantly KRT19-negative.

high expression of ALDH 3B1 (*ALDH3B1*) in LRECb were consistent with MaSC character. Similar to the situation in mouse and human, putative bovine MaSC (LRECb) appear to be ESR1-negative (Capuco et al., 2009; **Figures 7E,F**), and increased ALDH activity is consistent with MaSC/progenitor character (Douvillie et al., 2009; Martignani et al., 2010; Rauner and Barash, 2012). Increased abundance of *HNF4A*, *NR5A2*, *NES*, *TERF1*, *NUP153*, and *FNDC3B* mRNA and decreased abundance of X-chromosome inactivation factor (*XIST*) in LRECb are noteworthy (**Table 1** and **Table S1** in Supplementary Material). Hepatocyte nuclear factor (*HNF4A*) is a liver stem cell transcription factor (Battle et al., 2006; Delaforest et al., 2011), *NR5A2* is a pluripotency transcription factor analogous to OCT4 (Heng et al., 2010), Nestin (*NES*) is a neural stem cell marker (Wiese et al., 2004), and *TERF1* (Telomeric repeat binding factor 1) is a marker for human and

mouse embryonic stem cells (Ginis et al., 2004). *FNDC3B* has been characterized as a marker of proliferation and cell migration. The absence or very low abundance of *XIST* in LRECb is consistent with MaSC identity, as absence of *XIST* expression and low *XIST* expression have been associated with hematopoietic stem and progenitor cells, respectively (Savarese et al., 2006). Transcripts of several genes that are involved in epigenetic modification of chromatin were also enriched in LRECb. Relative to ECb, LRECb expressed a greater number of transcription regulators, zinc fingers, and nuclear transporters (e.g., *NUP153*, *IPO13*). Importin 13 (*IPO13*) is a nucleocytoplasmic transport protein, which may serve as a marker for corneal epithelial progenitor cells (Wang et al., 2009). Because elements of the nuclear pore complex and importin are frequently down-regulated following cell differentiation (Yasuhara et al., 2009), increased expression of

NUP153 and *IPO13* in LRECb suggests that LRECb are undifferentiated epithelial cells. Recent research by Sherley and colleagues was undertaken to discover biomarkers for distributed stem cells, based upon identification of genes that are tightly coupled to asymmetric self renewal of cells in culture (Noh et al., 2011). Among the genes identified by these researchers, expression of *EPHX1*, *MTBP*, *COL11A1*, and *ARHGAP* was increased in LRECb in the current experiment. Finally, expression of cytokeratin markers was consistent with expression by MaSC. The basal epithelial cells were KRT19-negative (**Figure 1E**), and transcriptome analysis indicated that *KRT5* was strongly down-regulated in LRECb, consistent with MaSC (Petersen and Polyak, 2010). Transcripts for fibroblast growth factors (*FGF1*, *FGF2*, *FGF10*), insulin-like growth factor-2 (*IGF2*) and follistatin (*FST*) were also enriched in LRECb. Overall, the gene expression profile of LRECb is consistent with MaSC character (**Tables 1 and 2**).

Further evidence in support of the stem cell nature of LRECb comes from biological pathway analysis of differentially expressed genes. Ingenuity Pathway Analysis of genes that were differentially expressed in LRECb and ECb revealed biological processes and networks that were highly significant. (Significance of a biologically relevant network of genes was expressed in IPA score, which was derived from *P*-value and indicates likelihood of the focused genes in a network being found together due to random chance. The IPA score is expressed as the negative log of the *P*-value.) The most significant networks associated with LRECb related to cellular growth and proliferation (**Figure 2A**, IPA score = 58), and cell cycle and post translational modification (**Figure 2B**, IPA score = 34). The network of cellular growth and proliferation (**Figure 2A**) contains a single module with *HNF4A*, up-regulated in LRECb, as the hub. Down-regulation of developmental genes like *SIX2* and *XIST* suggests that LRECb are undifferentiated cells. KEGG pathway analysis using DAVID (Huang et al., 2009) revealed that genes which were differentially expressed in LRECb vs. ECb reflected up-regulation of several pathways. These included the MAPK pathway (*FGF1*, *FGF2*, *FGF10*, *TAOK3*, *BRAF*, *ATF4*, *CREB*, *HSPA8*, *PDGFB*, *CDC25B*), a pathway involved in cellular growth and proliferation, and the WNT (*DVL2*, *PPP2R5E*, *SMAD4*) and TGF- β (*FST* and *SMAD4*) pathways, which are associated with stem cell renewal (Esmailpour and Huang, 2008; Mazumdar et al., 2010). In contrast to other members of the WNT pathway, *HOXA9* was strongly down-regulated in LRECb.

TRANSCRIPTOMES OF LRECe VS. ECe

Comparison of transcriptome profiles of LRECe and neighboring ECe identified 101 functionally annotated genes that were differentially expressed (Table S1 in Supplementary Material) and supports classification of LRECe as progenitor cells (**Table 1**). The most significant network associated with these genes was related to cancer (**Figure 3A**, IPA score = 51), followed by a network associated with DNA replication, recombination and repair (**Figure 3B**, IPA score = 36) that contained a *HNF4A* module. Conservation of the *HNF4A* module in LRECe and LRECb suggests a hierarchical similarity between LRECe and LRECb; although *HNF4A* transcripts were not significantly up-regulated in LRECe and genes involved































in this module differed between the two categories of LREC. Enriched expression of *NR5A2* and *FNDC3B* in both LRECe and LRECb (vs. ECe and ECb, respectively) provides another line of evidence for the similarity of LREC in basal and embedded epithelial layers. KEGG pathway analysis (DAVID) of transcripts that were up-regulated in LRECe vs. ECe identified up-regulation of the WNT pathway (*DVL3*, *ADCY6*, *CAMK2D*) and down-regulation of an inhibitor of the WNT pathway, (*CAMK2N1*).

TRANSCRIPTOMES OF LRECb VS. LRECe

To evaluate the relative characteristics of LRECb and LRECe, transcript profiles for these cells were compared. We identified 269 genes that were differentially expressed in LRECb vs. LRECe (Table S1 in Supplementary Material). Relatively high expression of stem cell markers, growth, and survival factors, DNA repair enzymes and low expression of apoptotic genes and differentiation markers supported greater “stemness” of LRECb vs. LRECe (**Tables 1–3**). The molecular profile of LRECb included increased abundance of transcripts for stem cell markers (*NR5A2*, *NES*, *THY1*), as well as cell survival and proliferation factors (*IGF2*, *FGF2*, *FGF10*, *HSPB6*, *LAMC1*, *CSF3*, *FST*, *IL33*, *MESDC2*, *AGT*). Additionally, enriched expression of cell adhesion molecules (*CADM3*, *NCAM1*, *AOC3*), and a number of cell surface markers (*ANXA6*, *CCR1*, *CCR4*, *CXCR4*, *DRD2*, *GNB4*, *GRB14*, *SAT2*, *SDPR*, *THY1/CD90*, *TRIB2*) were noted in LRECb. *THY1* has been used as a marker for hematopoietic progenitor cells (Goldschneider et al., 1978; Craig et al., 1993), mesenchymal stem cells (Gargett et al., 2009), and mammary cancer stem cells (Diehn et al., 2009). The LRECe displayed increased expression of *XIST*, splicing factor arginine/serine-rich 5 (*SFRS5*), *THAP* domain containing apoptosis associated protein 3 (*THAP3*), and calcium/calmodulin-dependent protein kinase II delta (*CAMK2D*). Increased expression of glucose metabolic enzymes, glucose phosphate isomerase (*GPI*), and UDP-glucose pyrophosphorylase 2 (*UGP2*) was also evident in LRECe vs. LRECb. KEGG pathway analysis (DAVID) revealed up-regulation of the Notch pathway in LRECb (*DVL2* an inhibitor of the pathway is down-regulated, *CAMK2D* and *MAML3* are up-regulated).






The most significant network associated with genes that were differentially expressed in LRECb vs. LRECe was related to tissue development, cell growth, and proliferation (**Figure 4A**, IPA score = 43). This network showed up-regulation in LRECb of *HIP1*, which may be required for differentiation or survival of somatic progenitors, and *TRIB2*, which modulates signal transduction pathways and may promote growth of mouse myeloid progenitors. This was followed by a network associated with tissue injury (**Figure 4B**, IPA score = 34), featuring up-regulation of a heat shock protein module in LRECb. The top three canonical pathways identified by IPA for genes that were preferentially expressed by LRECb (LRECb vs. LRECe) pertained to: the mitotic roles of polo-like kinases, cleavage, and polyadenylation of pre-mRNA, and chemokine signaling. Because polo-like kinases are key centrosome regulators and asymmetric localization of polo-kinase promotes asymmetric division of adult stem cells (Rusan and Peifer, 2007), the polo-like kinase pathway may be particularly noteworthy.

Table 1 | Attributes of LREC in prepubertal bovine mammary epithelium¹.





Gene	LRECb	LRECe	Description
STEMNESS MARKERS			
Transcription Factors			
<i>NR5A2</i>			Pluripotency transcription factor can substitute for OCT4 in production of induced pluripotent stem cells (iPSC)
	2.5	2.3	
<i>HNF4A</i>			Liver stem cell transcription factor
	2.0	1.2	
<i>SOX15</i>			Encodes a member of the SOX (SRY-related HMG-box) family of transcription factors involved in the regulation of embryonic development and in the determination of the cell fate
	2.4	1.3	
<i>ESR1</i>			LRECb are ESR1-negative (immunohistochemistry) and express low levels of ESR1 transcripts. ESR1 is a negative marker for MaSC in mouse and human. LRECe are ESR1-positive or negative
	-3.6	1.3	
Nuclear complex, pore, or transport			
<i>NUP153</i>			Nuclear basket protein, can cause chromatin modification, may be marker of proliferation, nuclear pore complex proteins are often down-regulated in differentiated cells
	7.1	-1.3	
<i>IP013</i>			Nucleocytoplasmic transport protein, importin 13, may serve as a marker for corneal stem and progenitor cells, nuclear transport plays a key role in stem cell lineage determination
	2.4	2.0	
<i>TERF1</i>			Telomeric repeat binding factor 1, marker for human and mouse embryonic stem cells
	3.7	1.0	
Cytoskeleton or membrane associated			
<i>THY1</i>			THY1/CD90 is up-regulated in MaSC enriched human mammary epithelial cultures [Expression of <i>THY1</i> in LRECb did not differ from that in ECb, but was 5.6-fold greater than that in LRECe]
	2.9	-2.3	
<i>HIP1</i>			May be required for differentiation or survival of somatic progenitors [Expression of <i>HIP1</i> in LRECb did not differ from that in ECb, but was 3.1-fold greater than that in LRECe]
	2.6	-1.3	
<i>NES</i>			Intermediate filament protein, expressed in stem cells: neural and hair follicle. [Expression of <i>NES</i> in LRECb did not differ from that in ECe, but was 2.6-fold greater than that in LRECe]
	-3.6	-1.3	
<i>FNDC3B</i>			Regulator of adipogenesis and cell proliferation, adhesion, spreading, and migration
	2.9	2.1	
Enzyme or Kinase			
<i>ALDH3B1</i>			Proposed MaSC marker
	2.0	1.3	
<i>TRIB2</i>			Modulates signal transduction pathways and may promote growth of mouse myeloid progenitors. [Expression of <i>TRIB2</i> in LRECb did not differ from that in ECb, but was 6.7-fold greater than that in LRECe]
	1.3	1.5	
Asymmetric cell renewal			
<i>EPHX1, MTBP, COL11A1</i>			Among proposed markers for somatic stem cells, based on genes that are tightly coupled to asymmetric self-renewal <i>in vitro</i> (Noh et al., 2011)
	2.3	1.0	
<i>ARHGAP</i>			
	6.3	1.6	

(Continued)

Table 1 | Continued

Gene	LRECb LRECe	Description
DIFFERENTIATION MARKERS		
Transcript Factor		
<i>XIST</i>	 -17.1 10.4	No to little expression of X-chromosome inactivation factor in LRECb, absence or low expression has been associated with hematopoietic stem and progenitor cells, respectively
<i>SIX2</i>	 -3.2 -1.1	Down-regulated in LRECb, is a homeobox protein, transcription factor. Like other members of this family, it may be involved in differentiation, involved in limb or eye development
<i>PAX6</i>	 -2.7 2.0	Down-regulated in LRECb, is a paired box protein transcription factor. Implicated role in organogenesis. Possible role of PAX6 in breast cancer and tumorigenesis has recently been identified
Intermediate Filaments		
<i>KRT5</i>	 -4.7 -2.3	Cytokeratin 5 transcripts are down-regulated in LRECb, consistent with KRT5-negative MaSC in mouse and human
<i>KRT19</i>	 -2.4 1.1	The basal epithelium is typically KRT 19 negative, consistent with KRT19-negative MaSC in mouse and human. [Expression based on quantitative immunohistochemistry]

¹The transcript abundance in LRECb and LRECe are expressed relative to that in respective control cells. Abundance that varies significantly in LREC and control cells is depicted graphically, with the fold change provided below the graphic. Fold change is provided even for those genes whose abundance did not differ between the LREC class and its control cells (designated by open bar).

-  Transcripts were up-regulated greater than threefold change relative to respective control.
-  Transcripts were up-regulated greater than two but less than threefold change relative to respective control.
-  Transcripts were down-regulated greater than twofold change relative to respective control.
-  Transcripts abundance did not differ from respective control.

TRANSCRIPTOMES OF ECb VS. ECe

Epithelial cells isolated from basal and embedded layers exhibited transcriptome profiles that were consistent with their location. Analysis identified 317 genes that were differentially expressed (Table S1 in Supplementary Material), 263 of which were functionally annotated. Among these, ECb expressed increased transcript levels for cell structural and motility genes, including actin (*ACTA2*), myosin (*MYH8*, *MYO6*, *MRCL3*), *SPTBN1* (actin cross linking scaffold protein), and *TSPAN31*. Transcripts for *JAG-1* (ligand of Notch pathway) and FST like 1 (*FSTL1*) were enriched in basal epithelium. The enriched expression of integrin- β 1 (*ITGB1*) within ECb was consistent with its use as a marker to isolate MaSC (Shackleton et al., 2006), most likely to enrich the sorted population for basal epithelial cells. Additionally, a number of heat shock proteins (*HSPA8*, *HSPA4*, *HSP90AB1*), peptidases (*USP4*, *USP16*, *USP25*, *PSMD14*, *MME*), ribosomal proteins, translational regulators, components of the ECM and its regulators (*collagens*, *MFAP5*, *FBN1*, *FSTL1*, *CHAD*, *ERBB2IP*, *SPARC*), and tumor suppressors [*MYCBP2*, and *MTSS1* (LOC788499)] were also up-regulated in ECb. However, transcripts of membrane transporters (*APIM1*, *APOE*, *AQP7*, *SLC13A3*, *SLC38A3*, *TMED3*, *CLCN3*) were more highly expressed in ECe than ECb. Thus, control cells harvested from basal and from embedded layers within the mammary epithelium possess different characteristics and appear to represent two distinct cell populations.

To better understand key biological processes occurring in basal and embedded epithelium, we utilized Ingenuity Pathway Analysis to generate gene networks and canonical pathways for genes that are differentially expressed between ECb and ECe. All identified networks (networks of endocrine system development and function, cancer, cell cycle, tissue development) were highly significant as measured by IPA score (ranges from 35 to 42). The identified network for endocrine development and function, lipid metabolism (**Figure 5A**) features an estrogen signaling module, peptidase, Ubiquitination, and ubiquitin modules. The identified network for cancer (**Figure 5B**) contains two heat shock protein modules. The canonical pathways identified by IPA analysis were protein ubiquitination, hypoxia signaling, and clathrin mediated endocytosis. Extrinsic growth factors and regulators, and hypoxia inducing factor have been identified as molecules prevalent in the stem cell niche (Li and Xie, 2005; Mazumdar et al., 2010), transcripts for these molecules are expressed in the basal epithelium (**Table 2; Figure 6**).

IMMUNOHISTOCHEMICAL AND REALTIME RT-PCR EVALUATION OF POTENTIAL NOVEL LRECb AND LRECe MARKERS

Genes that are highly expressed in LRECb and LRECe may provide novel markers for MaSC and progenitor cells. Those that were evaluated by immunohistochemistry were: *NR5A2*, *NUP153*, *FNDC3B*, and *HNF4A*. *NR5A2* is a pluripotency gene that aids in inducing somatic cells into pluripotency (iPSC; Heng et al., 2010).

Table 2 | Pathways in LREC and microenvironment of the basal mammary epithelium.

PATHWAY INVOLVEMENT WITH LREC	
Notch pathway	Up-regulation of Notch pathway in LRECb. Involvement of Notch pathway in LRECb along with pathways promoting expansion of MaSC is consistent with the need for balanced expansion of both MaSC and progenitor cell populations. Both populations must expand during ductal elongation
TGF- β signaling	Up-regulation in LRECb of components of TGF- β pathway associated with stem cell renewal
MAPK pathway	Up-regulation in LRECb, pathway involved in cell growth and proliferation
WNT pathway	Involvement of WNT pathway. Up-regulation of many components in LRECb. Presumed activity in LRECe > LRECb
MICROENVIRONMENT OF BASAL MAMMARY EPITHELIUM	
LRECb	Fibroblast growth factors (<i>FGF2</i> , <i>FGF10</i>), insulin-like growth factor-2 (<i>IGF2</i>), follistatin (<i>FST</i>), cytokines (<i>IL33</i> , <i>CSF3</i>), Wnt chaperone (<i>MESDC2</i>), heat inducible factors (<i>HIF1A</i> , <i>HSP</i>)
ECb	Production of factors by ECb that may impact LRECb: <i>JAG-1</i> (a ligand of Notch pathway) fibroblast growth factors (<i>FGF1</i> , <i>FGF2</i> , <i>FGF10</i>), insulin-like factor-2 (<i>IGF2</i>), follistatin like 1 (<i>FSTL1</i>), extracellular matrix and its regulators (collagens, <i>ITGB1</i> , <i>MFAP5</i> , <i>FBN1</i> , <i>FSTL1</i> , <i>CHAD</i> , <i>ERBB2IP</i> , <i>SPARC</i>), and tumor suppressors (<i>MTSS1</i> , Myc binding proteins)
Stromal cells	Gene expression by stromal cells was not assessed. However, these cells also contribute to the microenvironment of the basal epithelium

NUP153 is a nuclear basket protein that can cause chromatin modification (Vaquerizas et al., 2010), and *FNDC3B* is a regulator of adipogenesis and cell proliferation, adhesion, spreading, and migration (Nishizuka et al., 2009). *HNF4A* may serve as a stem cell regulator (Battle et al., 2006; Koh et al., 2010; Delaforest et al., 2011) and was identified as a key pathway component by IPA analysis of expression data for LRECb and LRECe. Transcripts for *NR5A2*, *NUP153*, *FNDC3B*, and *HNF4A* were more abundant in LRECb than in control cells, with a general expression pattern of LRECb > LRECe > EC).

Immunohistochemical analysis showed that 1–6% of epithelial cells expressed these potential markers. In agreement with transcript abundance, positive cells in the basal epithelium were more intensely stained than those in suprabasal locations. The abundance and localization of *NR5A2*, *NUP153*, *FNDC3B*, and *HNF4A*-positive cells (Figures 7A–D) were similar to that of LRECs. Co-localization studies showed that LREC expressed these markers. Surprisingly, expression of *FNDC3B* was not limited to the cytoplasmic compartment of the cell. Expression of *FNDC3B* was found to be cytoplasmic (arrows) and nuclear (arrowheads) and co-expressed with BrdU in approximately half of the LRECb

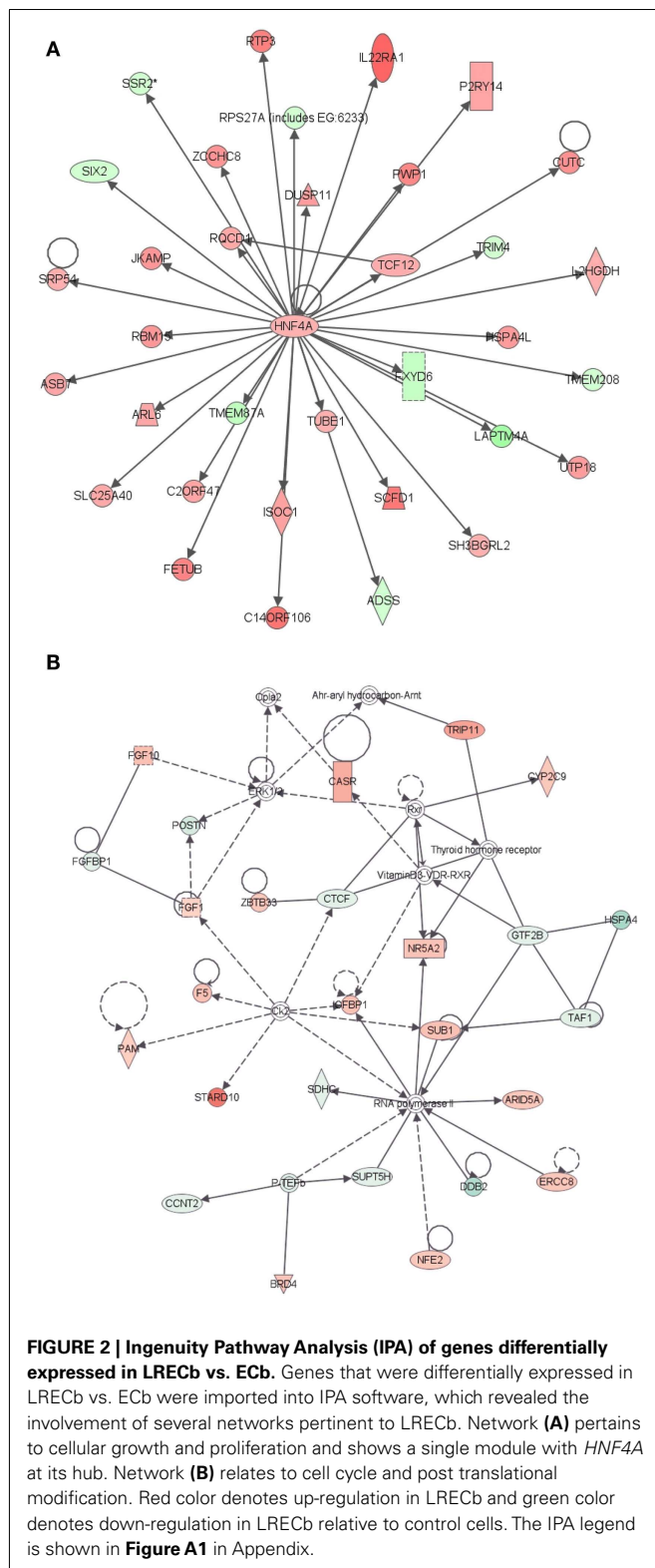
(Figure 7G), which is consistent with its possible utility as a marker for putative MaSC/progenitor cells. Co-localization studies also confirmed our previous finding (Capuco, 2007; Capuco et al., 2009) that LRECb are ESR1-negative and LRECe are composed of populations of ESR1-negative and ESR1-positive cells (Figures 7E,F). Because of their potential utility for cell sorting, we also identified transcripts that encoded surface proteins and were up-regulated in LRECb (*SAT2*, *CXCR4*, *SDPR*, *RTP3*, *CASR*, *GNB4*, and *DRD2*); however, we have not evaluated the suitability of these membrane markers. Preliminary immunohistochemistry results showed that *CXCR4* and *CASR* are expressed by a small number of epithelial cells.

Realtime RT-PCR was employed to confirm microarray results for expression of transcripts for novel LREC-derived markers (*NR5A2*, *NUP153*, *FNDC3B*) and the differentiation factor *XIST* at the transcriptome level. Patterns of expression were very similar for RT-PCR and microarray analysis (Figures 8A–C). Both analyses showed that expression of the potential MaSC/progenitor cell markers was increased in LRECb and, with the exception of *NUP153*, in LRECe vs. their respective controls. Expression of these markers was greater in LRECb vs. LRECe by microarray analysis, but *NR5A2* expression was not greater in LRECb vs. LRECe when assessed by realtime RT-PCR. Consistent with the undifferentiated state of putative MaSC, there was little to no expression of the differentiation factor *XIST* in LRECb, and there was lower expression of *XIST* in LRECb than in LRECe by both methodologies. Expression of *XIST* non-coding RNA was less in LRECe than in control cells as assessed by RT-PCR, but greater when assessed by microarray hybridization. Overall, the utility of microarray data for detecting LREC-derived markers for putative MaSC/progenitor cells was supported by realtime RT-PCR and by immunohistochemistry.

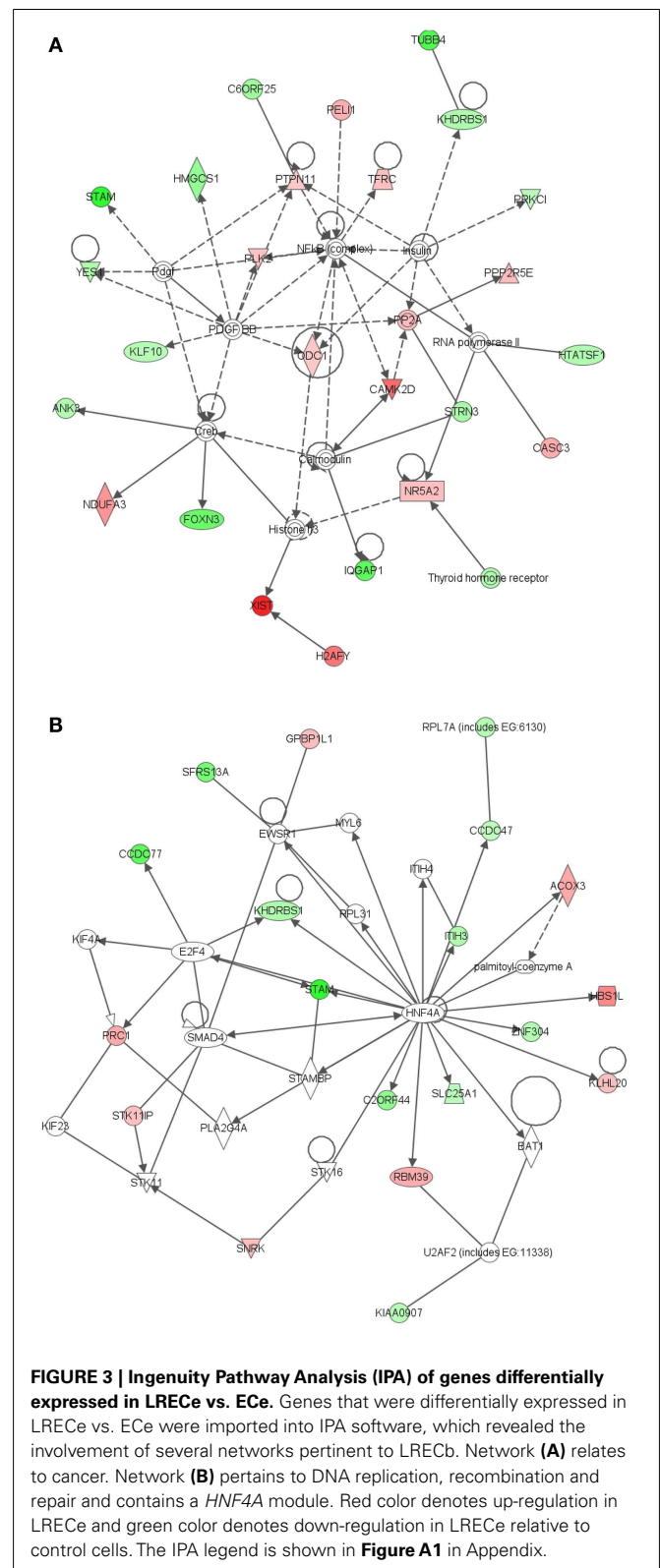
DISCUSSION

In this study, we employed the long-term retention of BrdU-labeled DNA to identify putative MaSC/progenitor cells during the period of ductal morphogenesis in the prepubertal mammary gland. However, it must be understood that retention of labeled DNA represents an integration of a cell's past proliferation and differentiation events and may not reflect that cell's current status. This is particularly relevant when assessing individual cells within a population, e.g., expression of lineage markers by LREC. Nonetheless, we hypothesized that LREC are enriched for MaSC/progenitor cells. In particular, we hypothesized that LRECb are enriched for MaSC and LRECe are enriched for progenitor cells.

When comparing gene expression in LREC and control cells it is important to consider the proliferative status of these cells. A difference in the proliferative status of the two populations may impose differences in gene expression between the populations that are reflective of their relative cell cycle activity rather than cell lineage. To determine the extent to which LREC proliferate during ductal morphogenesis in the prepubertal bovine mammary gland, we evaluated expression of nuclear proliferation antigens. In the present experiment, we found that approximately 13% of LREC and 15% of control cells in the present experiment expressed PCNA (data not shown). In previous studies, we evaluated the Ki-67 labeling index in calves at an equivalent stage



of mammary development to those in the present experiment and reported that 5.4% of LREC expressed Ki-67 (Capuco, 2007) and that 5–8% of total epithelial cells expressed Ki-67 (Capuco et al., 2004). Thus, the proliferation status of LREC, control cells



and the overall epithelial population appear to be similar and not likely to unduly influence interpretation of gene expression data.

Table 3 | Top 10 up-regulated and down-regulated genes in LRECb vs. LRECe.

Gene symbol	Bovine RefSeq ID	Gene ID	Gene annotation	Fold change
UP-REGULATED IN LRECb				
<i>LOC786372</i>	XM_001254067.2	297472984	PREDICTED: <i>Bos taurus</i> uncharacterized	14.3
<i>RPS7</i>	NM_001098874	402744265	<i>Bos taurus</i> ribosomal protein S7	12.5
<i>CFL1</i>	NM_001015655.1	62751776	<i>Bos taurus</i> cofilin 1 (non-muscle)	8.3
<i>RRN3</i>	NM_001192432	329112519	<i>Bos taurus</i> RRN3 RNA polymerase I transcription factor homolog (S. <i>cerevisiae</i>)	7.1
<i>COL1A2</i>	NM_174520.2	31341767	<i>Bos taurus</i> collagen, type I, alpha 2	7.1
<i>TRIB2</i>	NM_178317.3	58332431	<i>Bos taurus</i> tribbles homolog 2 (<i>Drosophila</i>)	6.7
<i>HSPB6</i>	NM_001076027.1	115496723	<i>Bos taurus</i> heat shock protein, alpha-crystallin-related, B6	6.7
<i>NUP153</i>	NM_001205754	329664715	<i>Bos taurus</i> nucleoporin 153kDa	6.3
<i>SLC10A5</i>	XM_001253131.2	194672838	PREDICTED: <i>Bos taurus</i> similar to solute carrier family 10 (sodium/bile acid cotransporter family), member 5	5.9
<i>RPL13</i>	NM_001015543	402692502	<i>Bos taurus</i> ribosomal protein L13	5.9
DOWN-REGULATED IN LRECb				
<i>XIST</i>	NR_001464.2	166706871	<i>Bos taurus</i> X (inactive)-specific transcript, non-coding RNA	−5.6
<i>C19H17orf49 (BAP18)</i>	NM_001038092.1	84000150	<i>Bos taurus</i> chromosome 17 open reading frame 49 ortholog	−3.9
<i>SFRS5</i>	NM_001098929.1	149643058	<i>Bos taurus</i> splicing factor, arginine/serine-rich 5	−3.7
<i>THAP3</i>	NM_001075347.1	115496479	<i>Bos taurus</i> THAP domain containing, apoptosis associated protein 3	−3.7
<i>STK11IP</i>	XM_593410	358411062	PREDICTED: <i>Bos taurus</i> similar to LKB1 interacting protein	−3.7
<i>UGP2</i>	NM_174212.2	110347575	<i>Bos taurus</i> UDP-glucose pyrophosphorylase 2	−3.4
<i>SLC3A2</i>	NM_001024488.2	164448599	<i>Bos taurus</i> solute carrier family 3 (activators of dibasic and neutral amino acid transport), member 2	−3.3
<i>CAMK2D</i>	NM_001046333.1	114052473	<i>Bos taurus</i> calcium/calmodulin-dependent protein kinase II delta	−3.3
<i>EIF4E3</i>	NM_0010102306.1	156121256	<i>Bos taurus</i> eukaryotic translation initiation factor 4E family member 3	−3.3
<i>ATXN7L3B</i>	NM_001078161.2	210147444	<i>Bos taurus</i> ataxin 7-like 3B	−3.0

To address our hypothesis that LRECb are enriched for MaSC and that LRECe are enriched for more committed progenitors, we performed transcriptome analyses on the four populations of bovine mammary epithelial cells obtained by laser microdissection of LREC and EC from basal and embedded layers of the epithelium. Microarray analysis was used to reveal gene signatures for the four categories of mammary epithelial cells: LRECb, LRECe, ECb, and ECe.

The ECb and ECe were distinguishable by the increased abundance, in basal cells, of transcripts for genes encoding structural and motility proteins, extracellular growth factors, extracellular matrix (ECM) proteins, and ECM regulators. Additionally, increased expression of transcripts for heat shock proteins, peptidases, ribosomal proteins, ubiquitins, proteins that provide interaction between the cell and the ECM (caveolin-1, integrin-beta-1), tumor suppressors, and epigenetic modifiers were also characteristic of ECb. Myoepithelial cells, present in the basal layer of mature mammary epithelium, may be a part of the stem cell niche and their paracrine factors may regulate the proliferation, polarity, and motility of mammary epithelial cells (Polyak and Hu, 2005). However, the precise nature of the ECb in a calf is uncertain. Expression of markers for myoepithelial cells in mammary tissue from prepubertal heifers is absent or expressed in a limited fashion (Capuco et al., 2002; Ballagh et al., 2008; Ellis et al., 2012; Safayi et al., 2012).

Transcriptome analysis of LRECb vs. ECb showed that LRECb possess characteristics consistent with those of MaSC (Tables 1 and 2). Our mRNA data indicated a reduced expression of *ESR1*

and increased expression of *ALDH3B1* in LRECb vs. ECb, and immunohistochemistry demonstrated a lack of detectable *ESR1* protein in LRECb. Previous studies have demonstrated that mouse (Sleeman et al., 2007) human (Anderson and Clarke, 2004) and putative bovine MaSC are *ESR1*-negative (Capuco et al., 2009). *ALDH1* activity has been used as a stem and progenitor cell marker in several tissues including blood, lung, prostate, pancreas, and breast (Duvillie et al., 2009). However, 17 isoforms of *ALDH* have been identified (Sladek, 2003) with different cellular and species expression patterns (Hess et al., 2004). *ALDH3B1* is expressed by bovine LRECb. Increased abundance of *HNF4A*, *NR5A2*, *TERF1*, *THY1*, *NUP153*, and *FNDC3B* mRNA and decreased abundance *XIST* transcripts (non-coding) in LRECb are noteworthy. *HNF4A* is a hepatic stem cell transcription factor whose associated network was highly up-regulated in LRECb, suggesting a key role in these cells. It is noteworthy that *HNF4A* has recently been implicated as a regulator of mesenchymal stem cells (Koh et al., 2010). Lack of expression or low expression of *XIST* has been associated with stem and progenitor cells, respectively, in hematopoietic tissue (Savarese et al., 2006). Subsequently, we evaluated four potentially novel protein markers for stem/progenitor cells (*NR5A2*, *NUP153*, *FNDC3B*, *HNF4A*) immunohistochemically and found protein expression profiles that were consistent with the observed transcript abundance in LRECb and LRECe. The number of cells expressing these markers was limited and staining intensity of the positive cells was greater for those located in the basal layer of the epithelium.

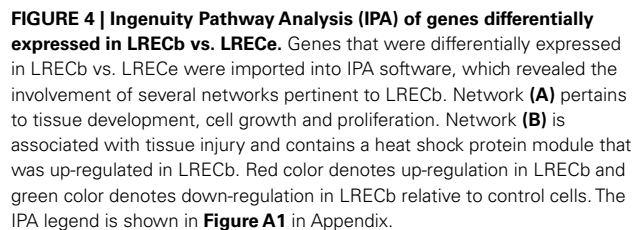
[illegible]

FIGURE 5 | Ingenuity Pathway Analysis (IPA) of genes differentially expressed in ECb vs. ECe. Genes that were differentially expressed in ECb vs. ECe were imported into IPA software, which revealed the involvement of several networks pertinent to ECb. Network **(A)** pertains to endocrine development and function, lipid metabolism. Network **(B)** is associated with cancer and contains two heat shock protein modules. Red color denotes up-regulation in LRECb and green color denotes down-regulation in LRECb relative to control cells. The IPA legend is shown in **Figure A1** in Appendix.

2006). *SAT2* is the target of DNA methyltransferase 1 (*DNMT1*) and an epigenetic modifier, whose methylation status may serve as a marker for cancer prognosis (Jackson et al., 2004). *CXCR4* is a receptor for the chemokine, stromal derived factor 1 (*SDF-1*; Kang et al., 2005). *SDF-1* is positively regulated by HIF1A, linking the SDF-CXCR4 axis to hypoxic stress. G-protein signaling

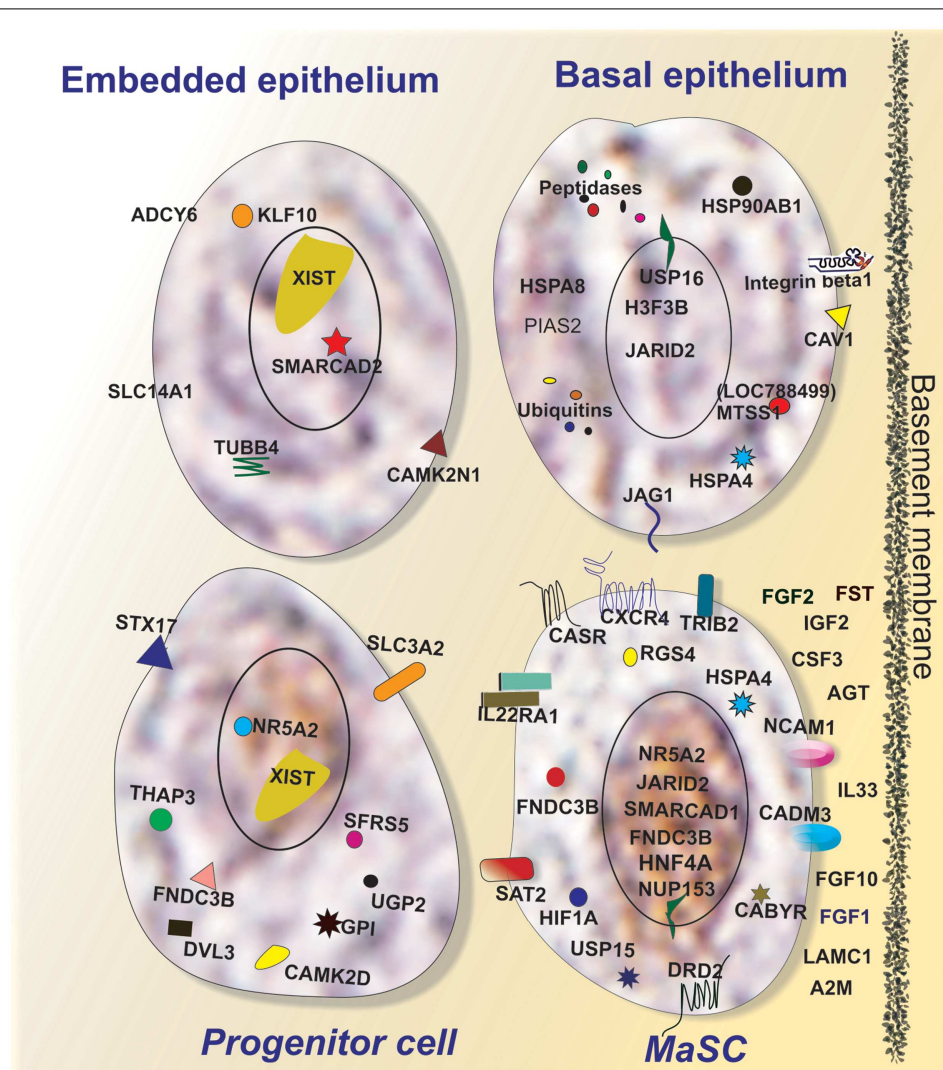


FIGURE 6 | Characteristics of putative MaSC (LRECb) and progenitor cells (LRECe). Putative MaSC (LRECb) are localized in the basal epithelium in a stem cell niche characterized as hypoxic but an environment enriched for extracellular growth factors, tumor suppressors for regulating MaSC function. LRECb exhibit enriched expressions of adhesion molecules and a variety of

potential MaSC biomarkers including HNF4A and the pluripotency marker, NR5A2. Putative progenitor cells (LRECe) also express NR5A2 but at a reduced level along with increased expression of differentiation factors including XIST. Key features of control cells (non-LREC) in the basal and embedded epithelial layers are also depicted.

proteins such as RGS4, which was up-regulated in LRECb, are negative regulators of the SDF-CXCR4 axis. The pertinence of the SDF-CXCR4 axis to stem cell regulation is the likelihood that mild hypoxic stress induces expansion of the MaSC population analogous to the expansion of breast cancer stem cells (Conley et al., 2012).

Up-regulation of growth factors such as fibroblast growth factors (*FGF1*, *FGF2*, *FGF10*), insulin-like factor-2 (*IGF2*), *FST*, laminin (*LAMC2*), platelet-derived growth factor beta (*PDGFB*), and plasminogen activator tissue (*PLAT*) in the basal epithelial layer is consistent with the possible function of these molecules as regulators of MaSC. The role of FGFs in mammary gland development and growth has been demonstrated (Mailleux et al., 2002; Sinowatz et al., 2006). Although our data do not provide

evidence for enhanced expression of receptors for these growth factors in LRECb, transcripts for many of these receptors were evident.

Further evidence in support of LRECb being a population of cells that is enriched for MaSC comes from biological pathway analysis of differentially expressed genes. A number of differentially expressed genes (LRECb vs. Ecb) were involved in MAPK, WNT, and TGF- β pathways. The MAPK pathway regulates cellular growth and proliferation. WNT and TGF- β pathways are both involved in mammary stem cell renewal. Down-regulation of TGF- β leads to a decline in MaSC number (Petersen and Polyak, 2010). A theme emerging from a variety of data is that stem cells exhibit characteristics of cells under stress (Covello et al., 2006; Mazumdar et al., 2010). An up-regulation of chaperones,

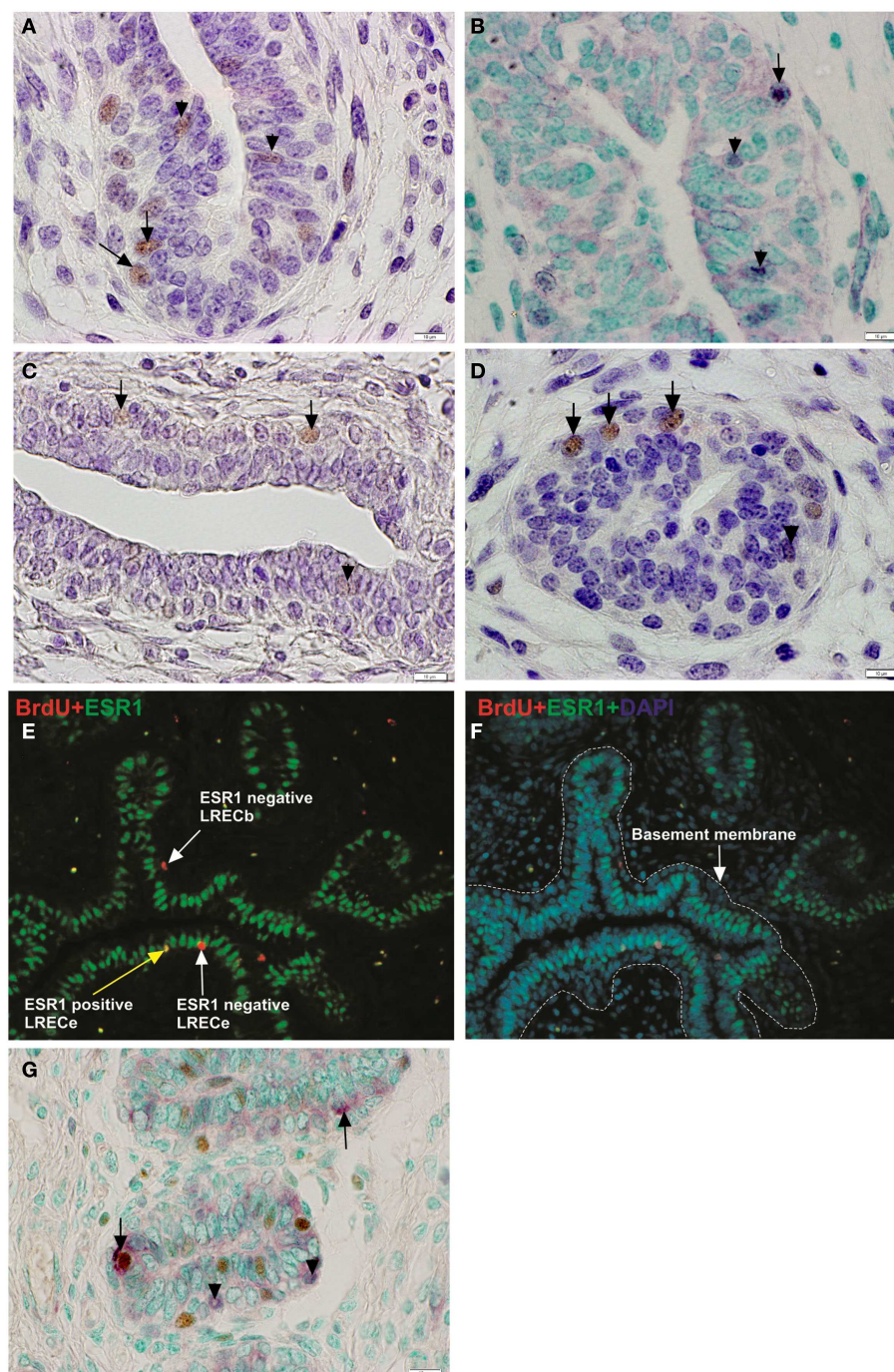
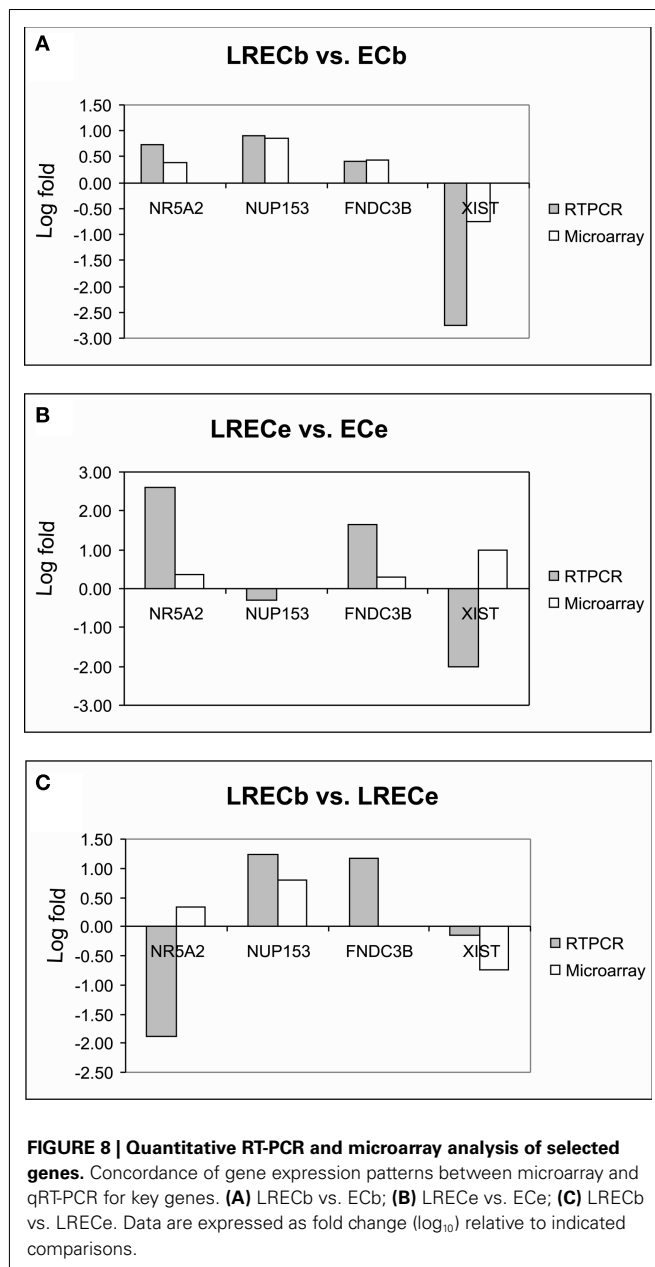


FIGURE 7 | Immunohistochemical localization of potential mammary stem/progenitor cell markers. (A–D) Consistent with transcriptome data, cells that were positive for these novel markers were more strongly labeled (brown nuclei) in the basal layer of mammary epithelium than in the embedded layers. Solid arrows designate labeled nuclei of basal cells and arrow heads designate labeled nuclei of embedded epithelial cells. NR5A2

(A), NUP153 (B), FNDC3B (C), HNF4A (D). (E,F) Immunofluorescent micrograph demonstrating the ESR1 status of LREC. LRECb were ESR1-negative and LRECe were a mixed population of ESR1-negative and ESR1-positive cells. (G) Co-localization of FNDC3B (purple) and BrdU. FNDC3B was expressed in the nucleus (arrow) and cytoplasm (arrowhead) of LRECb (brown). Scale bar, 10 μ m.

ubiquitin/proteasome, DNA repair, and chromatin remodeling in LRECb are consistent with this characteristic and support our hypothesis that LRECb are enriched for MaSC.

Comparison of the transcript profiles of LRECe with those of ECe and LRECb supports classification of LRECe as progenitor cells. As with LRECb, presence of an HNF4A network and BrdU



label retaining ability suggest that LRECe possess some stem cell attributes. However, up-regulation of metabolic enzymes and differentiation factors suggest that LRECe are more differentiated than LRECb. *XIST* is a non-coding RNA that inactivates one of the X-chromosomes in the early embryo and initiates gene repression and defines epigenetic transitions during development. Pluripotency genes (*NANOG*, *OCT4* and *SOX2*) cooperate to repress *XIST* (Navarro et al., 2008). Our mRNA data revealed low expression of *XIST* in LRECb and greater expression in LRECe and ECb, consistent with classification of LRECb as MaSC and LRECe as progenitor cells (Savarese et al., 2006). Finally, comparison of transcript abundance in LRECb vs. LRECe suggested up-regulation of the Notch pathway in LRECb, implying increased transduction of

Table 4 | Attributes of murine somatic stem cells and bovine mammary LREC¹.

Murine somatic cells	LRECb	LRECe
GEP, SiEP		
Wnt/ β -catenin signaling	+	
PI3K/AKT signaling		
TGF β signaling	+	+
IGF-1 signaling	IGF2	
JAK/Stat regulation		
Cell Cycle (G1/S) Checkpoint	+	
NF κ B signaling	+	+
GEP, HSCs, NSCs, ESCs		
Wnt/ β -catenin signaling	+	
PI3K/AKT signaling		
Interferon signaling	+	
PDGF signaling		
T cell receptor signaling		
Estrogen receptor signaling	+	+
GEps, SiEPs, HSCs, NSCs, ESCs		
Chemokine signaling	+	+
Integrin signaling		
SAPK/JNK signaling		
VEGF signaling		
IGF-1 signaling	IGF2	
B cell receptor signaling		
SiEPs, HSCs, NSCs, ESCs		
Wnt/ β -catenin signaling		
IL4 and IL6 signaling	IL2, IL12	
Alanine and aspartate metabolism		
P38 MAPK signaling	+	
Integrin signaling		

¹Properties of murine somatic stem cells, assessed by Gordon and colleagues (Giannakis et al., 2006), and the relevance of these properties to bovine mammary LRECb. GEPs, gastric epithelial precursors; SiEPs, small intestine epithelial precursors; HSCs, hematopoietic stem cells; NSCs, neural stem cells; ESCs, embryonic stem cells.

Notch signals in LRECb. The Notch pathway plays a critical role in cell fate determination of human mammary stem and progenitor cells (Dontu et al., 2004). In murine mammary gland, the Notch pathway constrains MaSC expansion and promotes proliferation and commitment to the luminal lineage (Bouras et al., 2008). Involvement of Notch signaling in putative MaSC (LRECb) along with pathways regulating stem cell expansion is consistent with the need to promote and balance the expansion of both MaSC and luminal epithelial cells during ductal mammarygenesis.

Using laser microdissection and RNA-sequencing, Gordon and colleagues evaluated the transcriptomes of progenitor cells and differentiated cells in the gastrointestinal tracts of mice, discerned characteristics of these precursors and compared their molecular properties with those of stem/progenitor cells in other organs (Stappenbeck et al., 2003; Giannakis et al., 2006). The use of laser microdissection was efficacious and led to the identification of characteristics that are shared among various stem cells. Many of

the molecular features of gastrointestinal and other adult stem cells that were identified are also evident in mammary LRECb (**Table 4**), supporting the hypothesis that the LRECb population is enriched for MaSC. Surprisingly, Gene Ontology-based analysis of transcripts that are differentially enriched in LREC and EC were inconsistent with the previously reported conclusion (Doherty et al., 2008) that stem cells exhibit increased expression of genes that are involved in nuclear function and RNA binding, while differentiated cells are enriched for expression of genes that are involved in extracellular space, signal transduction, and the plasma membrane (data not shown).

Our study provides supportive evidence that the stem cell niche lies in the basal layer of mammary epithelium. In this study, LREC and control cells were isolated from known locations within the mammary epithelium without previously destroying cellular microenvironments. However, LRECb were probably not in direct contact with the stroma, but were likely insulated by underlying cytoplasmic extensions from surrounding cells. The dissected LREC and control cells were adjacent or in close proximity to allow evaluations of potential cross-talk between putative MaSC and neighboring cells. Although potential signals were evident, additional research is necessary to elucidate such cross-talk. Furthermore, this analysis cannot account for signals that are derived from adjacent stromal cells, which were not interrogated. Microarray analyses of LRECb and LRECe identified features of LRECb that are reflective of MaSC residing in their stem cell niche. Distinct features of a stem cell niche, as discussed by Li and Xie (2005), are the presence of (1) cell adhesion molecules that provide anchorage for stem cells within the niche, (2) extrinsic factors within the niche that regulate stem cell behavior, and (3) factors that cause asymmetric cell division of the stem cell, that is upon cell division, one daughter cell is maintained in the niche as a stem cell (self renewal) and the other daughter cell leaves the niche to proliferate and differentiate. Recent studies also indicated that a stem cell niche elicits characteristics of hypoxic stress in stem cells, resulting in the induction of proteins of the family of hypoxia inducible transcription factors (HIF), such as *HIF1A* (up-regulated in LRECb), and targets *WNT*, *OCT4*, *IGF2* and Notch signaling molecules (Kaufman, 2010; Mazumdar et al., 2010). Mild hypoxia appeared to elicit expansion of mammary tumor stem cells via a mechanism mediated by *HIF1A* (Conley et al., 2012). In our study, we identified specific cell adhesion molecules, extrinsic growth factors and regulators, factors that promote asymmetric cell division, and hypoxia inducing factor as molecules that are prevalent in the stem cell niche of the basal epithelium (**Table 1; Figure 6**).

Finally, this research has identified molecular markers that are enriched in LREC. Transcripts encoding the nuclear proteins NR5A2, NUP153, FNDC3B, and HNF4A were identified as potential markers for MaSC, as were transcripts encoding surface

proteins SAT2, THY1/CD90, CXCR4, SDPR, RTP3, CASR, GNB4, and DRD2. To our knowledge, none of these proteins have been tested or utilized as MaSC markers. Enrichment for MaSC has been based upon sorting for multiple markers, as no single marker has proved particularly efficacious. The utility of these markers for identification and for sorting of MaSC remains to be evaluated.

CONCLUSION

Transcriptome analysis of LREC and mammary epithelial cell sub-populations has provided a framework for future studies of normal mammary epithelial cell development and homeostasis, and for the pathobiology of breast cancer. First, our data support the utility of long-term retention of DNA label as a means to identify an enriched population of progenitor cells. The data support the hypothesis that LRECs are enriched for MaSC and progenitor cells, with LRECb being enriched for progenitors with more stemness features (putative MaSC) and LRECe being enriched for more committed progenitors. Second, our data support the contention that the basal layer of the mammary epithelium provides for the MaSC niche. Lastly, we offer the first transcriptome profile of putative MaSC (LRECb) and progenitor cells (LRECe) excised from their *in situ* locations and we have identified potential novel biomarkers for these cells.

Insights into the biology of stem cells will be gained by further confirmation of candidate MaSC markers proposed by this study. Such confirmation requires an evaluation of the self renewal and differentiation potential of cells expressing these markers. Identification of appropriate biomarkers will provide a means to identify MaSC and will facilitate our understanding of MaSC functions in mammary development, homeostasis, and cancer. Specific cell surface markers will provide a means for future isolation of MaSC and investigations of their biology.

ACKNOWLEDGMENTS

This work was funded by CRIS no. 1265-3200-083-00D from the USDA Agricultural Research Service and by National Research Initiative Competitive Grant no. 2008-35206-18825 from the USDA National Institute of Food and Agriculture.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Cancer_Genetics/10.3389/fonc.2013.00021/abstract

Table S1 | Microsoft excel worksheet listing transcripts whose abundance differs between cell types. Comparisons are: LRECb vs. ECb, LRECe vs. ECE, LRECb vs. LRECe, and ECb vs. ECE.

Table S2 | Microsoft Excel worksheet listing genes (partial list) enriched in LRECb vs. control cells (ECb) and their general functional categories.

REFERENCES

- Anderson, E., and Clarke, R. B. (2004). Steroid receptors and cell cycle in normal mammary epithelium. *J. Mammary Gland. Biol. Neoplasia* 9, 3–13.
- Asselin-Labat, M. L., Shackleton, M., Stingl, J., Vaillant, F., Forrest, N. C., Eaves, C. J., et al. (2006). Steroid hormone receptor status of mouse mammary stem cells. *J. Natl. Cancer Inst.* 98, 1011–1014.
- Ballagh, K., Korn, N., Riggs, L., Pratt, S. L., Dessauge, F., Akers, R. M., et al. (2008). Hot topic: prepubertal ovariectomy alters the development of myoepithelial cells in the bovine mammary gland. *J. Dairy Sci.* 91, 2992–2995.
- Battle, M. A., Konopka, G., Parviz, F., Gaggi, A. L., Yang, C., Sladek, F. M., et al. (2006). Hepatocyte nuclear factor 4alpha orchestrates expression of cell adhesion proteins during the epithelial transformation of the developing liver. *Proc. Natl. Acad. Sci. U.S.A.* 103, 8419–8424.
- Bickenbach, J. R. (1981). Identification and behavior of label-retaining cells in oral mucosa and skin. *J. Dent. Res.* 60 Spec No C, 1611–1620.

- Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19, 185–193.
- Bouras, T., Pal, B., Vaillant, F., Harburg, G., Asselin-Labat, M. L., Oakes, S. R., et al. (2008). Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. *Cell Stem Cell* 3, 429–441.
- Briskin, C., and Duss, S. (2007). Stem cells and the stem cell niche in the breast: an integrated hormonal and developmental perspective. *Stem Cell. Rev.* 3, 147–156.
- Cairns, J. (1975). Mutation selection and the natural history of cancer. *Nature* 255, 197–200.
- Capuco, A. V. (2007). Identification of putative bovine mammary epithelial stem cells by their retention of labeled DNA strands. *Exp. Biol. Med. (Maywood)* 232, 1381–1390.
- Capuco, A. V., Dahl, G. E., Wood, D. L., Moallem, U., and Erdman, R. E. (2004). Effect of bovine somatotropin and rumen-undegradable protein on mammary growth of prepubertal dairy heifers and subsequent milk production. *J. Dairy Sci.* 87, 3762–3769.
- Capuco, A. V., Ellis, S., Wood, D. L., Akers, R. M., and Garrett, W. (2002). Postnatal mammary ductal growth: three-dimensional imaging of cell proliferation, effects of estrogen treatment, and expression of steroid receptors in prepubertal calves. *Tissue Cell* 34, 143–154.
- Capuco, A. V., Evock-Clover, C. M., Minuti, A., and Wood, D. L. (2009). In vivo expansion of the mammary stem/progenitor cell population by xanthosine infusion. *Exp. Biol. Med. (Maywood)* 234, 475–482.
- Chin, M. H., Mason, M. J., Xie, W., Volinia, S., Singer, M., Peterson, C., et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5, 111–123.
- Chin, M. H., Pellegrini, M., Plath, K., and Lowry, W. E. (2010). Molecular analyses of human induced pluripotent stem cells and embryonic stem cells. *Cell Stem Cell* 7, 263–269.
- Choudhary, R. K., Daniels, K. M., Evock-Clover, C. M., Garrett, W., and Capuco, A. V. (2010). Technical note: a rapid method for 5-bromo-2(-deoxyuridine (BrdU) immunostaining in bovine mammary cryosections that retains RNA quality. *J. Dairy Sci.* 93, 2574–2579.
- Conboy, M. J., Karasov, A. O., and Rando, T. A. (2007). High incidence of non-random template strand segregation and asymmetric fate determination in dividing stem cells and their progeny. *PLoS Biol.* 5:e102. doi:10.1371/journal.pbio.0050102
- Conley, S. J., Gheordunescu, E., Kakarala, P., Newman, B., Korkaya, H., Heath, A. N., et al. (2012). Antiangiogenic agents increase breast cancer stem cells via the generation of tumor hypoxia. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2784–2789.
- Covello, K. L., Kehler, J., Yu, H., Gordan, J. D., Arsham, A. M., Hu, C. J., et al. (2006). HIF-2 α regulates Oct-4: effects of hypoxia on stem cell function, embryonic development, and tumor growth. *Genes Dev.* 20, 557–570.
- Craig, W., Kay, R., Cutler, R. L., and Lansdorf, P. M. (1993). Expression of Thy-1 on human hematopoietic progenitor cells. *J. Exp. Med.* 177, 1331–1342.
- Delaforest, A., Nagaoka, M., Si-Tayeb, K., Noto, F. K., Konopka, G., Battle, M. A., et al. (2011). HNF4A is essential for specification of hepatic progenitors from human pluripotent stem cells. *Development* 138, 4143–4153.
- Diehn, M., Cho, R. W., Lobo, N. A., Kalisky, T., Dorie, M. J., Kulp, A. N., et al. (2009). Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature* 458, 780–783.
- Doherty, J. M., Geske, M. J., Stappenbeck, T. S., and Mills, J. C. (2008). Diverse adult stem cells share specific higher-order patterns of gene expression. *Stem Cells* 26, 2124–2130.
- Dontu, G., Abdallah, W. M., Foley, J. M., Jackson, K. W., Clarke, M. F., Kawamura, M. J., et al. (2003). In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev.* 17, 1253–1270.
- Dontu, G., Jackson, K. W., McNicholas, E., Kawamura, M. J., Abdallah, W. M., and Wicha, M. S. (2004). Role of Notch signaling in cell-fate determination of human mammary stem/progenitor cells. *Breast Cancer Res.* 6, R605–R615.
- Douville, J., Beaulieu, R., and Balicki, D. (2009). ALDH1 as a functional marker of cancer stem and progenitor cells. *Stem Cells Dev.* 18, 17–25.
- Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene Expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210.
- Ellis, S., Akers, R. M., Capuco, A. V., and Safayi, S. (2012). Bovine mammary epithelial cell lineages and parenchymal development. *J. Anim. Sci.* 90, 1666–1673.
- Esmailpour, T., and Huang, T. (2008). Advancement in mammary stem cell research. *J. Cancer Mol.* 4, 131–138.
- Gargett, C. E., Schwab, K. E., Zillwood, R. M., Nguyen, H. P., and Wu, D. (2009). Isolation and culture of epithelial progenitors and mesenchymal stem cells from human endometrium. *Biol. Reprod.* 80, 1136–1145.
- Giannakis, M., Stappenbeck, T. S., Mills, J. C., Leip, D. G., Lovett, M., Clifton, S. W., et al. (2006). Molecular properties of adult mouse gastric and intestinal epithelial progenitors in their niches. *J. Biol. Chem.* 281, 11292–11300.
- Ginis, I., Luo, Y., Miura, T., Thies, S., Brandenberger, R., Gerecht-Nir, S., et al. (2004). Differences between human and mouse embryonic stem cells. *Dev. Biol.* 269, 360–380.
- Goldschneider, I., Gordon, L. K., and Morris, R. J. (1978). Demonstration of Thy-1 antigen on pluripotent hemopoietic stem cells in the rat. *J. Exp. Med.* 148, 1351–1366.
- Heng, J. C., Feng, B., Han, J., Jiang, J., Kraus, P., Ng, J. H., et al. (2010). The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell* 6, 167–174.
- Hess, D. A., Meyerrose, T. E., Wirthlin, L., Craft, T. P., Herrbrich, P. E., Creer, M. H., et al. (2004). Functional characterization of highly purified human hematopoietic repopulating cells isolated according to aldehyde dehydrogenase activity. *Blood* 104, 1648–1655.
- Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.
- Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., et al. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Bio-statistics* 4, 249–264.
- Jackson, K., Yu, M. C., Arakawa, K., Fiala, E., Youn, B., Fiegel, H., et al. (2004). DNA hypomethylation is prevalent even in low-grade breast cancers. *Cancer Biol. Ther.* 3, 1225–1231.
- Kang, H., Watkins, G., Douglas-Jones, A., Mansel, R. E., and Jiang, W. G. (2005). The elevated level of CXCR4 is correlated with nodal metastasis of human breast cancer. *Breast* 14, 360–367.
- Kaufman, D. S. (2010). HIF hits Wnt in the stem cell niche. *Nat. Cell Biol.* 12, 926–927.
- Keeshan, K., He, Y., Wouters, B. J., Shestova, O., Xu, L., Sai, H., et al. (2006). Tribbles homolog 2 inactivates C/EBP α and causes acute myelogenous leukemia. *Cancer Cell* 10, 401–411.
- Kenney, N. J., Smith, G. H., Lawrence, E., Barrett, J. C., and Salomon, D. S. (2001). Identification of stem cell units in the terminal end bud and duct of the mouse mammary gland. *J. Biomed. Biotechnol.* 1, 133–143.
- Klein, A. M., and Simons, B. D. (2011). Universal patterns of stem cell fate in cycling adult tissues. *Development* 138, 3103–3111.
- Koh, W., Sheng, C. T., Tan, B., Lee, Q. Y., Kuznetsov, V., Kiang, L. S., et al. (2010). Analysis of deep sequencing microRNA expression profile from human embryonic stem cells derived mesenchymal stem cells reveals possible role of let-7 microRNA family in downstream targeting of hepatic nuclear factor 4 α . *BMC Genomics* 11(Suppl. 1):S6. doi:10.1186/1471-2164-11-S1-S6
- Lamarca, H. L., and Rosen, J. M. (2008). Hormones and mammary cell fate – what will I become when I grow up? *Endocrinology* 149, 4317–4321.
- Li, L., and Xie, T. (2005). Stem cell niche: structure and function. *Annu. Rev. Cell Dev. Biol.* 21, 605–631.
- Li, R. W., Meyer, M. J., Van Tassell, C. P., Sonstegard, T. S., Connor, E. E., Van Amburgh, M. E., et al. (2006). Identification of estrogen-responsive genes in the parenchyma and fat pad of the bovine mammary gland by microarray analysis. *Physiol. Genomics* 27, 42–53.
- Mailleux, A. A., Spencer-Dene, B., Dillon, C., Ndiaye, D., Savona-Baron, C., Itoh, N., et al. (2002). Role of FGF10/FGFR2b signaling during mammary gland development in the mouse embryo. *Development* 129, 53–60.
- Martignani, E., Eirew, P., Accornero, P., Eaves, C. J., and Baratta, M. (2010). Human milk protein production in xenografts of genetically engineered bovine mammary epithelial stem cells. *PLoS ONE* 5:e13372. doi:10.1371/journal.pone.0013372
- Mazumdar, J., O'Brien, W. T., Johnson, R. S., LaManna, J. C., Chavez, J. C., Klein, P. S., et al. (2010). O₂ regulates stem cells through Wnt/ β -catenin signalling. *Nat. Cell Biol.* 12, 1007–1013.
- Meyer, M. J., Capuco, A. V., Ross, D. A., Lintault, L. M., and Van

- Amburgh, M. E. (2006). Developmental and nutritional regulation of the prepubertal bovine mammary gland: II. Epithelial cell proliferation, parenchymal accretion rate, and allometric growth. *J. Dairy Sci.* 89, 4298–4304.
- Motyl, T., Bierla, J. B., Kozłowski, M., Gajewska, M., Gajkowska, B., and Koronkiewicz, M. (2011). Identification, quantification and transcriptional profile of potential stem cells in bovine mammary gland. *Livestock Sci.* 136, 136–149.
- Navarro, P., Chambers, I., Karwacki-Neisius, V., Chureau, C., Morey, C., Rougeulle, C., et al. (2008). Molecular coupling of Xist regulation and pluripotency. *Science* 321, 1693–1695.
- Niku, M., Ilmonen, L., Pessa-Morikawa, T., and Iivanainen, A. (2004). Limited contribution of circulating cells to the development and maintenance of nonhematopoietic bovine tissues. *Stem Cells* 22, 12–20.
- Nishizuka, M., Kishimoto, K., Kato, A., Ikawa, M., Okabe, M., Sato, R., et al. (2009). Disruption of the novel gene fad104 causes rapid postnatal death and attenuation of cell proliferation, adhesion, spreading and migration. *Exp. Cell Res.* 315, 809–819.
- Noh, M., Smith, J. L., Huh, Y. H., and Sherley, J. L. (2011). A resource for discovering specific and universal biomarkers for distributed stem cells. *PLoS ONE* 6:e22077. doi:10.1371/journal.pone.0022077
- Petersen, O. W., and Polyak, K. (2010). Stem cells in the human breast. *Cold Spring Harb. Perspect. Biol.* 2, a003160
- Polyak, K., and Hu, M. (2005). Do myoepithelial cells hold the key for breast tumor progression? *J. Mammary Gland Biol. Neoplasia* 10, 231–247.
- Potten, C. S., Hume, W. J., Reid, P., and Cairns, J. (1978). The segregation of DNA in epithelial stem cells. *Cell* 15, 899–906.
- Potten, C. S., Owen, G., and Booth, D. (2002). Intestinal stem cells protect their genome by selective segregation of template DNA strands. *J. Cell. Sci.* 115, 2381–2388.
- Rauner, G., and Barash, I. (2012). Cell hierarchy and lineage commitment in the bovine mammary gland. *PLoS ONE* 7:e30113. doi:10.1371/journal.pone.0030113
- Rusan, N. M., and Peifer, M. (2007). A role for a novel centrosome cycle in asymmetric cell division. *J. Cell Biol.* 177, 13–20.
- Safayi, S., Korn, N., Bertram, A., Akers, R. M., Capuco, A. V., Pratt, S. L., et al. (2012). Myoepithelial cell differentiation markers in prepubertal bovine mammary gland: effect of ovariectomy. *J. Dairy Sci.* 95, 2965–2976.
- Savarese, F., Flahndorfer, K., Jaenisch, R., Busslinger, M., and Wutz, A. (2006). Hematopoietic precursor cells transiently reestablish permissiveness for X inactivation. *Mol. Cell. Biol.* 26, 7167–7177.
- Shackleton, M., Vaillant, F., Simpson, K. J., Stingl, J., Smyth, G. K., Asselin-Labat, M. L., et al. (2006). Generation of a functional mammary gland from a single stem cell. *Nature* 439, 84–88.
- Shi, L., Jones, W. D., Jensen, R. V., Harris, S. C., Perkins, R. G., Goodsaid, F. M., et al. (2008). The balance of reproducibility, sensitivity, and specificity of lists of differentially expressed genes in microarray studies. *BMC Bioinformatics* 9(Suppl. 9):S10. doi:10.1186/1471-2105-9-S9-S10
- Shi, L., Reid, L. H., Jones, W. D., Shippey, R., Warrington, J. A., Baker, S. C., et al. (2006). The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat. Biotechnol.* 24, 1151–1161.
- Sinowatz, F., Schams, D., Habermann, F., Berisha, B., and Vermehren, M. (2006). Localization of fibroblast growth factor I (acid fibroblast growth factor) and its mRNA in the bovine mammary gland during mammaryogenesis, lactation and involution. *Anat. Histol. Embryol.* 35, 202–207.
- Sladek, N. E. (2003). Human aldehyde dehydrogenases: potential pathological, pharmacological, and toxicological impact. *J. Biochem. Mol. Toxicol.* 17, 7–23.
- Sleeman, K. E., Kendrick, H., Robertson, D., Isacke, C. M., Ashworth, A., and Smalley, M. J. (2007). Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland. *J. Cell Biol.* 176, 19–26.
- Smith, G. H. (2005). Label-retaining epithelial cells in mouse mammary gland divide asymmetrically and retain their template DNA strands. *Development* 132, 681–687.
- Smith, G. H., and Chepko, G. (2001). Mammary epithelial stem cells. *Microsc. Res. Tech.* 52, 190–203.
- Stappenbeck, T. S., Mills, J. C., and Gordon, J. I. (2003). Molecular features of adult mouse small intestinal epithelial progenitors. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1004–1009.
- Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., et al. (2006). Purification and unique properties of mammary epithelial stem cells. *Nature* 439, 993–997.
- Vaquerizas, J. M., Suyama, R., Kind, J., Miura, K., Luscombe, N. M., and Akhtar, A. (2010). Nuclear pore proteins nup153 and megator define transcriptionally active regions in the Drosophila genome. *PLoS Genet.* 6:e1000846. doi:10.1371/journal.pgen.1000846
- Wang, D., Peregrina, K., Dhima, E., Lin, E. Y., Mariadason, J. M., and Augenlicht, L. H. (2011). Paneth cell marker expression in intestinal villi and colon crypts characterizes dietary induced risk for mouse sporadic intestinal cancer. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10272–10277.
- Wang, H., Tao, T., Tang, J., Mao, Y. H., Li, W., Peng, J., et al. (2009). Importin 13 serves as a potential marker for corneal epithelial progenitor cells. *Stem Cells* 27, 2516–2526.
- Welm, B. E., Tepera, S. B., Venezia, T., Graubert, T. A., Rosen, J. M., and Goodell, M. A. (2002). Sca-1(pos) cells in the mouse mammary gland represent an enriched progenitor cell population. *Dev. Biol.* 245, 42–56.
- Wiese, C., Rolletschek, A., Kania, G., Blyszczuk, P., Tarasov, K. V., Tarasova, Y., et al. (2004). Nestin expression – a property of multi-lineage progenitor cells? *Cell. Mol. Life Sci.* 61, 2510–2522.
- Yasuhara, N., Oka, M., and Yoneda, Y. (2009). The role of the nuclear transport system in cell differentiation. *Semin. Cell Dev. Biol.* 20, 590–599.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 November 2012; paper pending published: 03 December 2012; accepted: 25 January 2013; published online: 15 February 2013.

Citation: Choudhary RK, Li RW, Evock-Clover CM and Capuco AV (2013) Comparison of the transcriptomes of long-term label retaining cells and control cells microdissected from mammary epithelium: an initial study to characterize potential stem/progenitor cells. *Front. Oncol.* 3:21. doi: 10.3389/fonc.2013.00021

This article was submitted to *Frontiers in Cancer Genetics*, a specialty of *Frontiers in Oncology*.

Copyright © 2013 Choudhary, Li, Evock-Clover and Capuco. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

APPENDIX

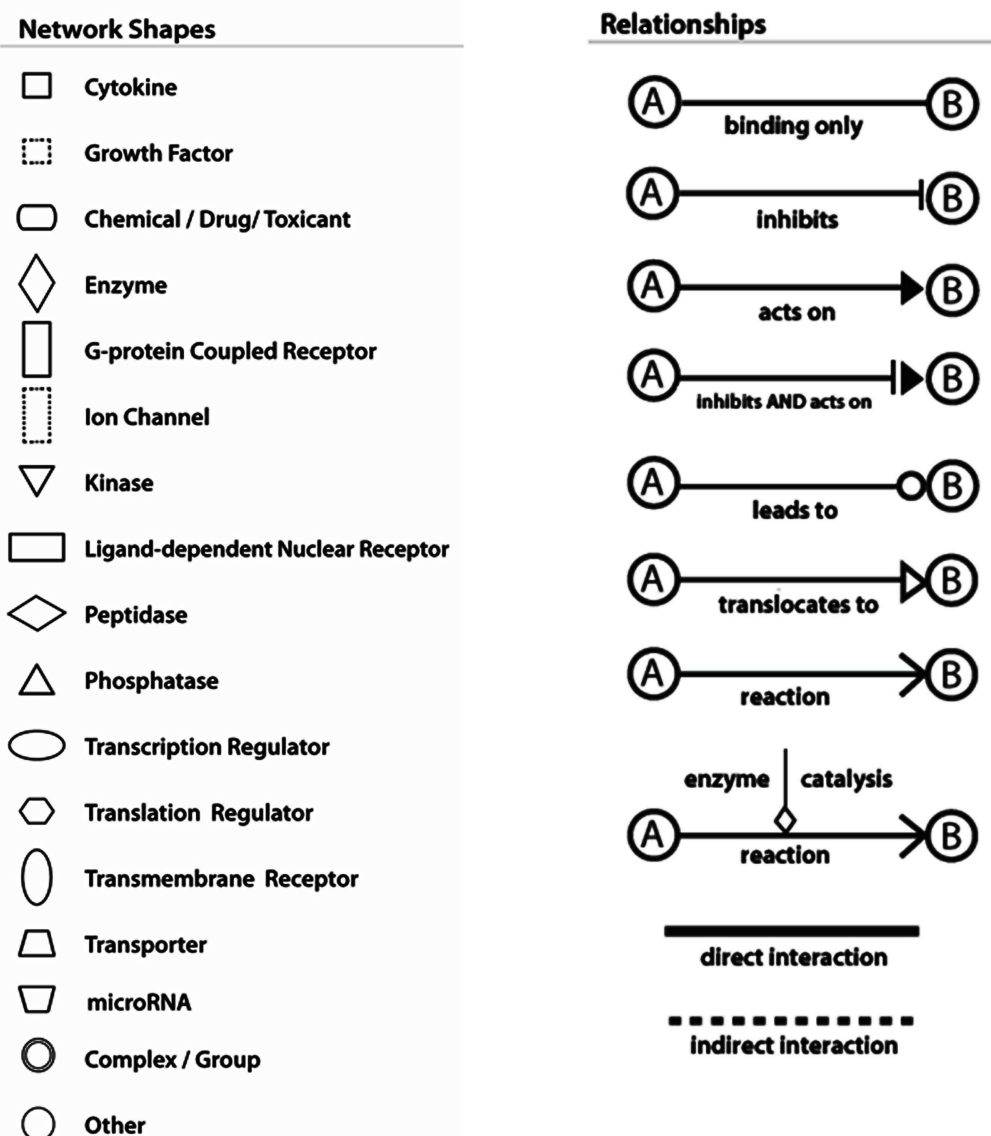


FIGURE A1 | JPG file defining the symbols and relationships depicted in IPA networks.



Asymmetric cell division and template DNA co-segregation in cancer stem cells

Sharon R. Pine* and Wenyu Liu

Department of Medicine, Robert Wood Johnson Medical School, Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA

Edited by:

James L. Sherley, The Adult Stem Cell Technology Center, LLC, USA

Reviewed by:

James L. Sherley, The Adult Stem Cell Technology Center, LLC, USA
Caterina A. M. La Porta, University of Milan, Italy

*Correspondence:

Sharon R. Pine, Rutgers Cancer Institute of New Jersey, 195 Little Albany Street, New Brunswick, NJ 08903, USA
e-mail: pinesr@cinj.rutgers.edu

During tissue homeostasis, normal stem cells self-renew and repopulate the diverse cell types found within the tissue via a series of carefully controlled symmetric and asymmetric cell divisions (ACDs). The notion that solid tumors comprise a subset of cancer stem cells (CSCs) with dysregulated self-renewal and excessive symmetric cell divisions has led to numerous studies aimed to elucidate the mechanisms regulating ACD under steady-state conditions, during stem-cell expansion and in cancer. In this perspective, we focus on a type of asymmetry that can be established during ACD, called non-random co-segregation of template DNA, which has been identified across numerous species, cell types, and cancers. We discuss the role of p53 loss in maintaining self-renewal in both normal and malignant cells. We then review our current knowledge of the mechanisms underlying co-segregation of template DNA strands and the stem-cell pathways associated with it in normal and CSCs.

Keywords: asymmetric cell division, immortal DNA, co-segregation of template DNA, p53, cancer stem cells, Wnt, Notch

INTRODUCTION

TISSUE HOMEOSTASIS

The continual maintenance of a reservoir of tissue-specific stem cells affords an organism with the ability to generate all the differentiated cells needed for tissue homeostasis and repair throughout its lifespan. During homeostasis or repair, a stem cell can undergo asymmetric cell division (ACD) to simultaneously generate daughters with differing cell fates. One daughter remains a stem cell and the other gives rise to differentiated progeny that carry out the functions of the mature tissue. When a stem cell asymmetrically divides, it does so by actively segregating one or more intrinsic cell fate-determining constituents, or by polarizing the cell cortex and aligning the mitotic spindle so that the two daughter cells are exposed to differing external stimuli directing their cell fate potential (1, 2). In certain circumstances, such as during development or after stem cell depletion from excessive injury, normal stem cells can also be exponentially expanded through a series of symmetric divisions. Thus, both asymmetric and symmetric cell divisions can lead to stem cell self-renewal.

ASYMMETRIC CELL DIVISION IN CANCER

It is necessary to maintain a tightly controlled balance between symmetric and asymmetric stem cell divisions in order to preserve an optimal number of stem cells within a tissue or organ. When the balance is shifted to favor excessive symmetric self-renewing divisions, it can lead to a hyperplastic state and cancer development [reviewed in Ref. (3)]. The tumor itself, however, can be viewed as an abnormal organ in which the mature tumor cells are seeded in a hierarchical fashion by a stem cell-like population, called cancer stem cells (CSCs). Unlike normal stem cells, CSCs have lost the ability to control their mode of cell division, resulting in continual

excessive symmetric cell divisions and consequential uncontrolled tumor growth (Figure 1).

A debated question in CSC biology is whether neoplastic transformation emerges from normal stem/progenitor cells or from more differentiated cells that commandeer stem cell properties during or after the oncogenic process. Since normal stem cells inherently possess many of the properties that CSCs exploit, it seems more likely that a tumor would arise from a normal stem cell. Lineage tracing experiments and genetically engineered mouse models in which oncogenic events were restricted to specific cell types have confirmed that transformation can indeed arise from the normal stem cell population (4). However, several reports have demonstrated that certain progenitor and differentiated cells can also be transformed [reviewed in Ref. (5)]. Thus, tumors comprise a hierarchically organized cell population seeded by CSCs regardless of the cell of origin. Notwithstanding, uncovering the mechanisms regulating normal stem cell self-renewing divisions could provide insights into how those mechanisms are disrupted in CSC biology, and lead to pharmacological strategies to deplete the CSC pool altogether.

Another intriguing feature in stem cell biology is cellular plasticity. It was long believed that tissues were hierarchically organized in which the tissue-specific stem cells reside at the apex and differentiation could only occur in one direction. However, recent *in vivo* lineage tracing experiments have proved that differentiated cells within a tissue can replenish the stem cell population (6, 7), especially after an event that induces stem-cell depletion. Our new understanding of normal cellular plasticity lends credence to the argument that cancer cells can dedifferentiate into CSCs. A normal stem cell or CSC that arises from a more differentiated cell would theoretically repossess the ability to symmetrically

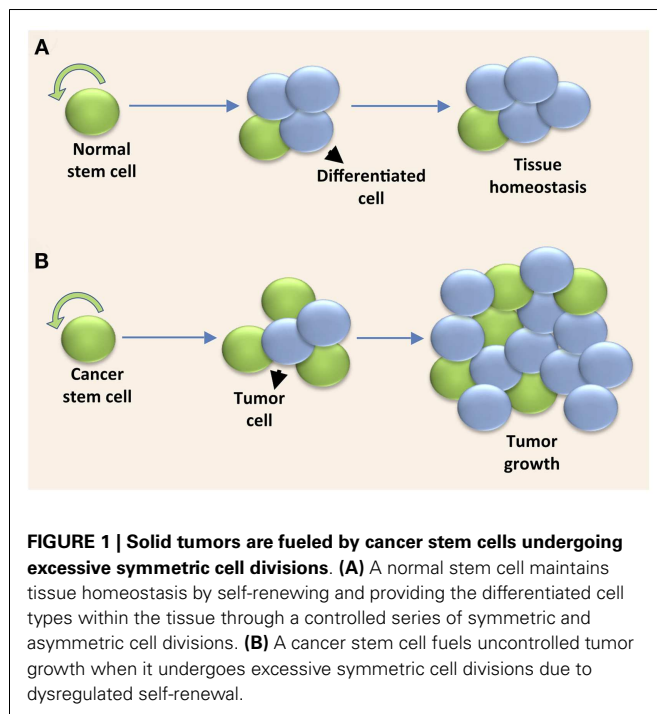


FIGURE 1 | Solid tumors are fueled by cancer stem cells undergoing excessive symmetric cell divisions. (A) A normal stem cell maintains tissue homeostasis by self-renewing and providing the differentiated cell types within the tissue through a controlled series of symmetric and asymmetric cell divisions. **(B)** A cancer stem cell fuels uncontrolled tumor growth when it undergoes excessive symmetric cell divisions due to dysregulated self-renewal.

or asymmetrically divide. How cellular plasticity might affect genomic integrity, proliferative lifespan of the resultant stem cell, or organismal aging is unknown.

CO-SEGREGATION OF TEMPLATE DNA STRANDS

One of the most intriguing types of asymmetry during ACD is the active segregation of the “older” template DNA strands specifically to one daughter cell and the shunting of the newly synthesized DNA strands to the other. As elegantly recounted by K. Gordon Lark in the research topic on stem cell genetic fidelity (8), the discovery of template DNA strand co-segregation during cell division resulted from a series of experiments in the mid to late 1960s (9). Even though it has been observed across numerous species and tissue types, the notion of template DNA co-segregation has been debated extensively because of a failure to observe it in some tissue stem cells, which might have been caused in some cases by the differing methods utilized to detect it [reviewed in Ref. (10)]. The initial observations by Lark led to theories explaining why cells would actively sort their DNA into old and new strands during cell division. One explanation, called the “immortal DNA strand hypothesis,” states that stem cells prevent cancer-causing replication-induced DNA mutations by exclusively co-segregating their older, “immortal” DNA strands to the long-lived daughter stem cells and passing their newly synthesized DNA off to the daughter cells that give rise to differentiated cells (11).

An alternative though not mutually exclusive hypothesis explains that the older template DNA strands harbor epigenetic marks that differ from those on the newer DNA strands. The differing epigenetic tags direct differential gene expression after cell division, leading to differing fates of the daughter cells. If the older versus newer sets of sister chromatids harbor differing epigenetic marks, it is likely that the tags are added during DNA synthesis,

though this has not been definitively tested (12). The most convincing evidence supporting this hypothesis is from Amar Klar’s group, in which they demonstrated that specific sister chromatids are segregated during ACD in *Caenorhabditis elegans* and mice (13–15), which was based from their studies performed on the fission yeast *Schizosaccharomyces pombe*. Yeast sister cells developmentally differed by inheriting sister chromatids that were differentiated by epigenetic differences [reviewed in Ref. (16)]. While neither the immortal strand hypothesis nor epigenetic changes have been proven to be the driving functional consequence of co-segregation of template DNA strands in multicellular organisms; preservation of genomic integrity and gene expression patterns prior to stem cell division would both be beneficial to stem cell and organismal survival.

TEMPLATE DNA STRAND CO-SEGREGATION IN CANCER

In addition to normal tissues, we and others have found evidence for template DNA strand co-segregation in cell lines and short-term cultures of human tumors from numerous cancer types (17–22). It is unknown if the retention of template DNA co-segregation in cancer cells is a passive, albeit dysregulated, remnant of hierarchical organization in normal tissues, or if it offers a survival advantage. The mutation theory of the immortal DNA strand hypothesis is far more relevant in normal stem cells compared to malignant cells, because genomic instability and consequential tumor progression would be favored for tumor cell survival. But epigenetic regulation plays a central role in tumor progression (23), so priming the daughter cells with an epigenetic signature prior to cell division could potentially accelerate tumor progression.

We reported that the frequency of template DNA co-segregation in lung and breast cancer cell lines decreases when the microenvironment favors symmetric self-renewal (17, 18), and that the template DNA strands are inherited by the daughter cell with more CSC-like qualities (17). Therefore, it is likely that the mechanism underlying the decision to self-renew is an active process at least partly under the control of extrinsic factors. Elucidation of the molecular mechanisms driving template DNA strand co-segregation in cancer would provide important clues for why and how CSCs undergo excessive self-renewal.

p53 IN ASYMMETRIC CELL DIVISION AND TEMPLATE DNA STRAND CO-SEGREGATION

LOSS OF p53 DURING STEM CELL SELF-RENEWAL

p53 is the most extensively studied tumor suppressor gene and is often referred to as the “guardian of the genome” (24). The best understood functions of p53 are in the orchestration of cellular responses to different stress stimuli through the induction of cell cycle arrest, DNA repair, senescence, and apoptosis. Recently, p53 was found to function in the maintenance of a stem cell state. Work that has emerged from reprogramming somatic cells into pluripotent cells, referred to as induced pluripotent stem cells (iPSCs), has confirmed the role of the p53 pathway in the production of this type of stem cell. Loss or knockdown of p53 not only increases reprogramming efficiency, it also accelerates the kinetics of reprogramming [reviewed in Ref. (25)], though it is not yet clear exactly how loss of p53 contributes to improved reprogramming efficiency. There has

been mounting evidence that loss of p53 contributes to normal stem cell self-renewal, especially after DNA damage. An excellent example supporting this comes from work on mammary epithelial and hematopoietic stem cells. Pier Pelicci's group demonstrated that unlike progenitor and differentiated cells that activate p53 in response to DNA damage, p53 is not activated in normal stem cells. Intriguingly, p53-independent upregulation of p21 in normal stem cells after DNA damage inhibits p53 activity and shifts the cell divisions from asymmetric to symmetric self-renewal (26). Given the central role of p53 in normal stem cell self-renewal, and the conjecture that cancer is a disease of excessive self-renewal, it stands to reason that p53 mutations may induce excessive self-renewal of CSCs. Loss of p53 resulted in the expansion of pre-malignant mammary stem cells through increased symmetric self-renewing divisions, and when p53 was reactivated, there was a reduction in the CSC pool due to restoration of ACDs (27). Thus p53 likely plays a direct and central role in regulating the switch between asymmetric and symmetric cell divisions of both normal stem cells and CSCs.

p53 IN ASYMMETRIC CELL DIVISION

Sherley's group demonstrated that, in cell lines derived from murine embryonic fibroblasts and mammary epithelial cells, expression of wildtype p53 at physiological levels shifted the balance from symmetric to asymmetric stem cell kinetics. In this system, asymmetric self-renewal was defined as a cell giving rise to a stem-like cycling cell and a non-dividing daughter cell (28). Expression of p53 also altered co-segregation of template DNA. Under conditions that favored symmetric self-renewal in which p53 was not expressed, chromosomes were randomly segregated during cell division. In contrast, expression of p53, which favored asymmetric self-renewal, induced a shift toward increased co-segregation of template DNA strands (28). Thus, p53 induces asymmetric self-renewal of adult stem cells not only at a functional but also at a genomic level. These data support the notion that asymmetric self-renewal and template DNA strand co-segregation are at least sometimes coupled in stem cells.

THE ROLE OF DNA DAMAGE IN ASYMMETRIC CELL DIVISION

After a tissue damaging event, a large number of stem cells would be needed to quickly respond, expand, and replace damaged cells. However, cells could be most vulnerable to DNA replication errors during hyperproliferation and tissue regeneration. The fact that the genome would actively symmetrically segregate its damaged DNA to both daughter stem cells during rapid expansion seems counter-intuitive. According to the "immortal" DNA strand hypothesis, stem cells avoid passing their nascent DNA strands, wrought with potential DNA replication errors, off to its daughter stem cell while co-segregating its older and presumably undamaged DNA to its daughter stem cell. Why then would stem cells symmetrically self-renew and risk passing replication errors off to the daughter stem cells during a time of rapid cell cycling? An assumption of the "immortal" DNA strand hypothesis is that stem cell DNA segregation is under normal steady-state conditions. Perhaps when faced with a DNA damaging event and expansion of the stem cell pool is necessary to regenerate lost cells, the symmetrically self-renewing stem cells employ an as yet

undiscovered DNA damage response pathway unique to repairing the stem cell genome. Based on the work from Pelicci's group as described above, the pathway is likely to be independent of p53. Furthermore, mouse embryonic stem cells (ESCs) undergo repair double strand breaks much faster than differentiated cells, primarily using homologous recombination [reviewed in Ref. (29)]. Given that reduction of p53 supports replication-associated homologous recombination, it is possible that this DNA repair pathway prevails in symmetrically dividing somatic stem cells, though this would need to be tested. A stem cell-associated DNA repair pathway would also complement Cairn's original observations in which he found a mathematical discrepancy between predicted DNA mutation rates in human tissue cells and human cancer incidence (11). Furthermore, it was proposed that chemical modifications or DNA mutations in the "immortal" stem cell genome contributes to aging (30). Together with co-segregation of template DNA strands, such a DNA repair mechanism during symmetric self-renewal would protect the stem cell genome from not only DNA replication errors but also DNA damage, and possibly even the cellular aging process.

p53 ISOFORMS

We would be remiss to discuss the central role of p53 in stem cell ACD without mentioning p53 isoforms. The human TP53 gene encodes at least 13 isoforms that are the result of alternative splicing or alternative promoters (31). Two of the best studied human p53 isoforms functionally interact with full-length p53. Isoform $\Delta 133p53$ is an N-terminally truncated isoform that inhibits full-length p53 in a dominant-negative manner. Isoform p53 β is a C-terminally truncated isoform that cooperates with full-length p53 (32). Curtis Harris' group, with whom Sharon R. Pine had her post-doctoral fellowship training, demonstrated that p53 isoform switching can modulate the youthfulness of a cell. They showed that p53 β represses and $\Delta 133p53$ increases the replicative lifespan of normal human fibroblasts cultured *in vitro*. Furthermore, p53 isoform switching was associated with tumor progression in colon cancer (33), consistent with the modes of functional interaction between these p53 isoforms and full-length p53. They further demonstrated that p53 isoform switching directly modulates cellular replicative lifespan, senescence, and activation of CD8⁺ T lymphocytes *in vivo* (34). An untested question to date is whether p53 isoforms are associated with ACD. This could have important implications in cancer biology because p53 isoform expression patterns and activity would theoretically depend on the location of the TP53 mutation. We did not observe a correlation between p53 mutation or deletion status and frequency of template DNA asymmetric segregation in a panel of breast cancer cell lines (18). It would be intriguing to test if specific p53 isoforms are involved in the regulation ACD or co-segregation of template DNA strands.

MOLECULAR MECHANISMS OF TEMPLATE DNA STRAND CO-SEGREGATION

Template DNA strand co-segregation has been observed in stem cells across numerous tissue types, but the underlying molecular mechanism(s) regulating this process has remained elusive. In order to decipher how template chromosomes are asymmetrically segregated during cell division, one must identify what "marks"

the newer and older DNA strands, what cellular machinery recognizes those marks and finally, what process actively recruits the chromosomes bearing those marks to one daughter cell. The Tajbakhsh group speculated that the mother centrosome might anchor certain sister chromatids to the polarized cortex (35). How the kinetochore might relay information to the centrosome and cell cortex via the mitotic spindle to mediate asymmetric chromosome segregation is still unsolved.

Much of the effort to identify the molecular mechanisms responsible for a cell's decision to co-segregate its template DNA strands has been pioneered by the Sherley group. Rambhatla et al. demonstrated that template DNA co-segregation was induced by p53 and required the down-regulation of an enzyme essential for guanine nucleotide biosynthesis, IMP dehydrogenase (36). Perhaps the mechanism by which p53 regulates template DNA strand co-segregation is linked to the concentration of guanine nucleotides, although it could also be linked to one or more of many other cellular functions regulated by guanine nucleotides, such as signal transduction, glycoprotein synthesis, or energy transfer. It was not clear if down-regulation of IMP dehydrogenases was sufficient for p53-induced template DNA co-segregation, or if one or more of the numerous p53-regulated genes were also key modulators of template DNA strand co-segregation.

Sherley's group later discovered that the template DNA strands harbor more of the histone H2A variant H2A.Z that is "uncloaked", meaning that it is more detectable by immunofluorescence (37). They more recently reported that the template DNA strands had a higher content of 5-hydroxymethylcytosine (5hmC) than the newer DNA strands, which is an intermediate during DNA methylation (38). Although 5hmC, or a unique protein complex that binds 5hmC in stem cells, could possibly be the elusive "mark" of template DNA strands, differences in DNA demethylation between the two daughter cells of an asymmetric stem cell division are also consistent with the idea that the template DNA strand co-segregation dictates cell fate via a differential regulation of gene expression in the two daughter cells. Klar's group discovered additional clues that could underlie the mechanisms of template DNA co-segregation. They found that in *S. pombe* and *Schizosaccharomyces japonicas* yeast that a DNA strand-specific epigenetic imprint at the mating locus initiates a recombination event, which is required for cellular differentiation (16, 39). In their system, an inherent chirality of the double-helical structure of DNA was needed to achieve ACD. They later provided evidence for this during mouse development (13). An intriguing question is whether a similar epigenetic mechanism is conserved on a global level in multicellular organisms to dictate different cell fate potentials of the daughter cells.

Wnt PATHWAY

Additional clues about the mechanisms driving template DNA strand co-segregation have emerged from studies of stem-cell signaling pathways that are either associated with or actively affect the balance of ACD and template DNA strand co-segregation. Wnt signaling is a key stem cell signaling pathway implicated in self-renewal of both normal stem cells and CSCs, as well as epithelial-mesenchymal transition in cancer (40). Inhibition of

Wnt signaling decreased the frequency of template DNA co-segregation in gastrointestinal cancer cell lines (21). This work was challenged in a separate study in which random segregation of template DNA in colorectal CSCs was induced by Wnt signaling (41). It would be interesting to test if Wnt inhibition or activation reduces template DNA co-segregation by increasing symmetric divisions of non-stem cells, or increasing symmetric self-renewing divisions within the stem cell population, respectively. Although it was uncertain whether Wnt signaling played a direct role in regulating template DNA co-segregation, or if one of the many pathways modulated by Wnt/ β -catenin signaling were responsible, these studies demonstrated that factors in addition to p53 or IMP dehydrogenase may directly modulate co-segregation of template DNA, and particularly, genes within signaling pathways that direct stem cell fate.

NOTCH PATHWAY

The most extensively studied pathway associated with template DNA co-segregation is the Notch signaling pathway. Notch is a developmental and adult stem/progenitor cell transcription factor that regulates self-renewal and differentiation in a highly cell-type and context-specific manner (42). Notch signaling is dysregulated across numerous types of cancer and plays a key role in regulating CSC self-renewal (43). The Notch pathway directly participates in ACD across several cell types in lower organisms such as *Drosophila* and *C. elegans*, as well as in mice and human beings (1). For example, in *Drosophila* neuroblasts the Notch inhibitor Numb is inherited by the daughter cell that is destined to undergo differentiation (1, 44). When Numb is mutated, the neuroblasts hyperproliferate and form a tumor-like phenotype (45, 46). Notch dysregulation is also an active participant during the development of cancer in mouse models (47–49), though whether the role of Notch in carcinogenesis involves shifting the mode of CSC divisions toward excessive symmetric self-renewal has not been established.

Notch signaling has been linked to template DNA co-segregation. One example has been demonstrated in muscle satellite stem cells. Satellite cells self-renew and provide the generation of myoblasts that are needed for skeletal muscle homeostasis and repair. Muscle satellite stem cells co-segregate their template DNA strands (50, 51) and the daughter stem cells that inherit the template DNA preferentially inherit Numb (50). In colon CSCs, an increase in Notch mRNA levels caused by knockdown of miR34a increases symmetric self-renewing divisions and decreases ACDs (52). It is still unclear whether Notch signaling is a cellular constituent that is asymmetrically segregated in parallel with but independent from template strand co-segregation, or if Notch signaling is a regulator or effector of template DNA co-segregation.

CONCLUSION

It has been nearly 50 years since the first discovery of template DNA co-segregation. Though it has been observed across numerous species, normal tissue, and cancer, we have still only scratched the surface toward elucidating the purpose for conserving template DNA strands in one daughter cell and the mechanisms regulating the process. Consistent with the notion that the immortal DNA strand hypothesis does not fully capture why stem cells asymmetrically segregate their template DNA, there has been

increasing evidence that the mother cell controls gene expression in the daughter cells through asymmetric segregation of epigenetic marks. With recent advances in genomics, chromosome-imaging technologies, and genetically modified model organisms [reviewed in Ref. (10)], the outlook for success in further answering these daunting questions is bright. Elucidating why and how template DNA strands are co-segregated is a fundamental aim in basic and translational science that spans many disciplines, and the implications could be profound. If we can manipulate self-renewal of stem cells, this could result in lucrative applications for directing self-renewal and differentiation for regenerative medicine, reversing degenerative diseases, as well as therapeutic interventions of cancer.

ACKNOWLEDGMENTS

This work was supported by the NIH, under Grant K22CA140719 (Sharon R. Pine).

REFERENCES

- Knoblich JA. Mechanisms of asymmetric stem cell division. *Cell* (2008) **132**(4):583–97. doi:10.1016/j.cell.2008.02.007
- Williams SE, Fuchs E. Oriented divisions, fate decisions. *Curr Opin Cell Biol* (2013) **25**(6):749–58. doi:10.1016/j.ceb.2013.08.003
- Gonczy P. Mechanisms of asymmetric cell division: flies and worms pave the way. *Nat Rev Mol Cell Biol* (2008) **9**(5):355–66. doi:10.1038/nrm2388
- Blanpain C. Tracing the cellular origin of cancer. *Nat Cell Biol* (2013) **15**(2):126–34. doi:10.1038/ncb2657
- Visvader JE. Cells of origin in cancer. *Nature* (2011) **469**(7330):314–22. doi:10.1038/nature09781
- Stange DE, Koo BK, Huch M, Sibbel G, Basak O, Lyubimova A, et al. Differentiated Troy+ chief cells act as reserve stem cells to generate all lineages of the stomach epithelium. *Cell* (2013) **155**(2):357–68. doi:10.1016/j.cell.2013.09.008
- Tata PR, Mou H, Pardo-Saganta A, Zhao R, Prabhu M, Law BM, et al. Dedifferentiation of committed epithelial cells into stem cells in vivo. *Nature* (2013) **503**(7475):218–23. doi:10.1038/nature12777
- Lark KG. Discovering non-random segregation of sister chromatids: the naive treatment of a premature discovery. *Front Oncol* (2012) **2**:211. doi:10.3389/fonc.2012.00211
- Lark KG, Consigli RA, Minocha HC. Segregation of sister chromatids in mammalian cells. *Science* (1966) **154**(3753):1202–5. doi:10.1126/science.154.3753.1202
- Yennek S, Tajbakhsh S. DNA asymmetry and cell fate regulation in stem cells. *Semin Cell Dev Biol* (2013) **24**(8–9):627–42. doi:10.1016/j.semcdb.2013.05.008
- Cairns J. Mutation selection and the natural history of cancer. *Nature* (1975) **255**(5505):197–200. doi:10.1038/255197a0
- Evano B, Tajbakhsh S. Sorting DNA with asymmetry: a new player in gene regulation? *Chromosome Res* (2013) **21**(3):225–42. doi:10.1007/s10577-013-9359-7
- Sauer S, Klar AJS. Left-right symmetry breaking in mice by left-right dynein may occur via a biased chromatid segregation mechanism, without directly involving the nodal gene. *Front Oncol* (2012) **2**:166. doi:10.3389/fonc.2012.00166
- Armakolas A, Klar AJ. Left-right dynein motor implicated in selective chromatid segregation in mouse cells. *Science* (2007) **315**(5808):100–1. doi:10.1126/science.1129429
- Armakolas A, Klar AJ. Cell type regulates selective segregation of mouse chromosome 7 DNA strands in mitosis. *Science* (2006) **311**(5764):1146–9. doi:10.1126/science.1120519
- Klar AJ. Lessons learned from studies of fission yeast mating-type switching and silencing. *Annu Rev Genet* (2007) **41**:213–36. doi:10.1146/annurev.genet.39.073103.094316
- Pine SR, Ryan BM, Varticovski L, Robles AI, Harris CC. Microenvironmental modulation of asymmetric cell division in human lung cancer cells. *Proc Natl Acad Sci U S A* (2010) **107**(5):2195–200. doi:10.1073/pnas.0909390107
- Liu W, Jegathan G, Amiri S, Morgan KM, Ryan BM, Pine SR. Asymmetric segregation of template DNA strands in basal-like human breast cancer cell lines. *Mol Cancer* (2013) **12**(1):139. doi:10.1186/1476-4598-12-139
- Hari D, Xin HW, Jaiswal K, Wiegand G, Kim BK, Ambe C, et al. Isolation of live label-retaining cells and cells undergoing asymmetric cell division via nonrandom chromosomal cosegregation from human cancers. *Stem Cells Dev* (2011) **20**(10):1649–58. doi:10.1089/scd.2010.0455
- Xue Z, Yan H, Li J, Liang S, Cai X, Chen X, et al. Identification of cancer stem cells in vincristine preconditioned SGC7901 gastric cancer cell line. *J Cell Biochem* (2012) **113**(1):302–12. doi:10.1002/jcb.23356
- Xin HW, Ambe CM, Ray S, Kim BK, Koizumi T, Wiegand GW, et al. Wnt and the cancer niche: paracrine interactions with gastrointestinal cancer cells undergoing asymmetric cell division. *J Cancer* (2013) **4**(6):447–57. doi:10.7150/jca.6896
- Varghese S, Whipple R, Martin SS, Alexander HR. Multipotent cancer stem cells derived from human malignant peritoneal mesothelioma promote tumorigenesis. *PLoS One* (2012) **7**(12):e52825. doi:10.1371/journal.pone.0052825
- Sharma S, Kelly TK, Jones PA. Epigenetics in cancer. *Carcinogenesis* (2010) **31**(1):27–36. doi:10.1093/carcin/bgp220
- Levine AJ, Oren M. The first 30 years of p53: growing ever more complex. *Nat Rev Cancer* (2009) **9**(10):749–58. doi:10.1038/nrc2723
- Bonizzi G, Cicalese A, Insinga A, Pelicci PG. The emerging role of p53 in stem cells. *Trends Mol Med* (2012) **18**(1):6–12. doi:10.1016/j.molmed.2011.08.002
- Insinga A, Cicalese A, Faretta M, Gallo B, Albano L, Ronzoni S, et al. DNA damage in stem cells activates p21, inhibits p53, and induces symmetric self-renewing divisions. *Proc Natl Acad Sci U S A* (2013) **110**(10):3931–6. doi:10.1073/pnas.1213394110
- Cicalese A, Bonizzi G, Pasi CE, Faretta M, Ronzoni S, Giulini B, et al. The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. *Cell* (2009) **138**(6):1083–95. doi:10.1016/j.cell.2009.06.048
- Merok JR, Lansita JA, Tunstead JR, Sherley JL. Cosegregation of chromosomes containing immortal DNA strands in cells that cycle with asymmetric stem cell kinetics. *Cancer Res* (2002) **62**(23):6791–5.
- Tichy ED. Mechanisms maintaining genomic integrity in embryonic stem cells and induced pluripotent stem cells. *Exp Biol Med (Maywood)* (2011) **236**(9):987–96. doi:10.1258/ebm.2011.011107
- Sherley JL. A new mechanism for aging: chemical “age spots” in immortal DNA strands in distributed stem cells. *Breast Dis* (2008) **29**:37–46.
- Khoury MP, Bourdon JC. The isoforms of the p53 protein. *Cold Spring Harb Perspect Biol* (2010) **2**(3):a000927. doi:10.1101/cshperspect.a000927
- Bourdon JC, Fernandes K, Murray-Zmijewski F, Liu G, Diot A, Xirodimas DP, et al. p53 isoforms can regulate p53 transcriptional activity. *Genes Dev* (2005) **19**(18):2122–37. doi:10.1101/gad.1339905
- Fujita K, Mondal AM, Horikawa I, Nguyen GH, Kumamoto K, Sohn JJ, et al. p53 isoforms Delta133p53 and p53beta are endogenous regulators of replicative cellular senescence. *Nat Cell Biol* (2009) **11**(9):1135–42. doi:10.1038/ncb1928
- Mondal AM, Horikawa I, Pine SR, Fujita K, Morgan KM, Vera E, et al. p53 isoforms regulate aging- and tumor-associated replicative senescence in T lymphocytes. *J Clin Invest* (2013) **123**(12):5247–57. doi:10.1172/JCI70355
- Tajbakhsh S, Gonzalez C. Biased segregation of DNA and centrosomes: moving together or drifting apart? *Nat Rev Mol Cell Biol* (2009) **10**(11):804–10. doi:10.1038/nrm2784
- Rambhatla L, Ram-Mohan S, Cheng JJ, Sherley JL. Immortal DNA strand cosegregation requires p53/IMPDH-dependent asymmetric self-renewal associated with adult stem cells. *Cancer Res* (2005) **65**(8):3155–61. doi:10.1158/0008-5472.CAN-04-3161
- Huh YH, Sherley JL. Molecular cloaking of H2A.Z on mortal DNA chromosomes during nonrandom segregation. *Stem Cells* (2011) **29**(10):1620–7. doi:10.1002/stem.707
- Huh YH, Cohen J, Sherley JL. Higher 5-hydroxymethylcytosine identifies immortal DNA strand chromosomes in asymmetrically self-renewing distributed stem cells. *Proc Natl Acad Sci U S A* (2013) **110**(42):16862–7. doi:10.1073/pnas.1310323110
- Yu C, Bonaduce MJ, Klar AJ. Defining the epigenetic mechanism of asymmetric cell division of Schizosaccharomyces japonicus yeast. *Genetics* (2013) **193**(1):85–94. doi:10.1534/genetics.112.146233
- Holland JD, Klaus A, Garratt AN, Birchmeier W. Wnt signaling in stem and cancer stem cells. *Curr Opin Cell Biol* (2013) **25**(2):254–64. doi:10.1016/j.ceb.2013.01.004

41. Hwang WL, Jiang JK, Yang SH, Huang TS, Lan HY, Teng HW, et al. MicroRNA-146a directs the symmetric division of Snail-dominant colorectal cancer stem cells. *Nat Cell Biol* (2014) **16**(3):268–80. doi:10.1038/ncb2910
42. Koch U, Lehal R, Radtke F. Stem cells living with a Notch. *Development* (2013) **140**(4):689–704. doi:10.1242/dev.080614
43. Capaccione KM, Pine SR. The Notch signaling pathway as a mediator of tumor survival. *Carcinogenesis* (2013) **34**(7):1420–30. doi:10.1093/carcin/bgt127
44. Siller KH, Doe CQ. Spindle orientation during asymmetric cell division. *Nat Cell Biol* (2009) **11**(4):365–74. doi:10.1038/ncb0409-365
45. Lee CY, Andersen RO, Cabernard C, Manning L, Tran KD, Lanskey MJ, et al. *Drosophila* aurora-A kinase inhibits neuroblast self-renewal by regulating aPKC/Numb cortical polarity and spindle orientation. *Genes Dev* (2006) **20**(24):3464–74. doi:10.1101/gad.1489406
46. Wang H, Somers GW, Bashirullah A, Heberlein U, Yu F, Chia W. Aurora-A acts as a tumor suppressor and regulates self-renewal of *Drosophila* neuroblasts. *Genes Dev* (2006) **20**(24):3453–63. doi:10.1101/gad.1487506
47. Allen TD, Rodriguez EM, Jones KD, Bishop JM. Activated Notch1 induces lung adenomas in mice and cooperates with Myc in the generation of lung adenocarcinoma. *Cancer Res* (2011) **71**(18):6010–8. doi:10.1158/0008-5472.CAN-11-0595
48. Shih HP, Kopp JL, Sandhu M, Dubois CL, Seymour PA, Grapin-Botton A, et al. A Notch-dependent molecular circuitry initiates pancreatic endocrine and ductal cell differentiation. *Development* (2012) **139**(14):2488–99. doi:10.1242/dev.078634
49. Licciulli S, Avila JL, Hanlon L, Troutman S, Cesaroni M, Kota S, et al. Notch1 is required for Kras-induced lung adenocarcinoma and controls tumor cell survival via p53. *Cancer Res* (2013) **43**:5974–84. doi:10.1158/0008-5472.CAN-13-1384
50. Shinin V, Gayraud-Morel B, Gomes D, Tajbakhsh S. Asymmetric division and cosegregation of template DNA strands in adult muscle satellite cells. *Nat Cell Biol* (2006) **8**(7):677–87. doi:10.1038/ncb1425
51. Conboy MJ, Karasov AO, Rando TA. High incidence of non-random template strand segregation and asymmetric fate determination in dividing stem cells and their progeny. *PLoS Biol* (2007) **5**(5):e102. doi:10.1371/journal.pbio.0050102
52. Bu P, Chen KY, Chen JH, Wang L, Walters J, Shin YJ, et al. A microRNA miR-34a-regulated bimodal switch targets Notch in colon cancer stem cells. *Cell Stem Cell* (2013) **12**(5):602–15. doi:10.1016/j.stem.2013.03.002

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 April 2014; accepted: 06 August 2014; published online: 21 August 2014.
Citation: Pine SR and Liu W (2014) Asymmetric cell division and template DNA co-segregation in cancer stem cells. Front. Oncol. 4:226. doi: 10.3389/fonc.2014.00226
This article was submitted to Cancer Genetics, a section of the journal Frontiers in Oncology.

Copyright © 2014 Pine and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Direct measurements of human colon crypt stem cell niche genetic fidelity: the role of chance in non-Darwinian mutation selection

Haeyoun Kang¹ and Darryl Shibata^{2*}

¹ Department of Pathology, CHA Bundang Medical Center, CHA University, Seongnam-si, South Korea

² Department of Pathology, University of Southern California Keck School of Medicine, Los Angeles, CA, USA

Edited by:

James L. Sherley, Boston Biomedical Research Institute, USA

Reviewed by:

Jian-Jun Wei, Northwestern University, USA

Allon Moshe Klein, Harvard Medical School, USA

*Correspondence:

Darryl Shibata, Department of Pathology, University of Southern California Keck School of Medicine, 1441 Eastlake Avenue, Los Angeles, CA 90033, USA
e-mail: dshibata@usc.edu

Perfect human stem cell genetic fidelity would prevent aging and cancer. However, perfection would be difficult to achieve, and aging is universal and cancers common. A hypothesis is that because mutations are inevitable over a human lifetime, downstream mechanisms have evolved to manage the deleterious effects of beneficial and lethal mutations. In the colon, a crypt stem cell architecture reduces the number of mitotic cells at risk for mutation accumulation, and multiple niche stem cells ensure that a lethal mutation within any single stem cell does not lead to crypt death. In addition, the architecture of the colon crypt stem cell niche may harness probability or chance to randomly discard many beneficial mutations that might lead to cancer. An analysis of somatic chromosome copy number alterations (CNAs) reveals a lack of perfect fidelity in individual normal human crypts, with age-related increases and higher frequencies in ulcerative colitis, a proliferative, inflammatory disease. The age-related increase in somatic CNAs appears consistent with relatively normal replication error and cell division rates. Surprisingly, and similar to point mutations in cancer genomes, the types of crypt mutations were more consistent with random fixation rather than selection. In theory, a simple “non-Darwinian” way to nullify selection is to reduce the size of the reproducing population. Fates are more determined by chance rather than selection in very small populations, and therefore selection may be minimized within small crypt niches. The desired effect is that many beneficial mutations that might lead to cancer are randomly lost by drift rather than fixed by selection. The subdivision of the colon into multiple very small stem cell niches may trade Darwinian evolution for non-Darwinian somatic cell evolution, capitulating to aging but reducing cancer risks.

Keywords: stem cell niche, non-Darwinian, neutral drift, human, replication errors, aging, Muller's ratchet

INTRODUCTION

The mechanisms responsible for human stem cell genetic fidelity are difficult to directly study because of the impracticality of experimental manipulations. However, the numbers and types of somatic mutations accumulated during a lifetime are end measures of stem cell genetic fidelity. It should be possible to infer from mutations the various mechanisms that may help limit both the accumulation and phenotypic consequences of human somatic mutations. Mutations can lead to cancer and aging, and replication errors are major potential sources of somatic alterations. The colon is highly mitotic organ, with most epithelial cells replaced weekly (1, 2). The colon is subdivided into millions of small clonal units called crypts. Human colon crypts are small (~2,000 cells) glands maintained by multiple stem cells and a stem cell hierarchy (Figure 1). A stem cell hierarchy helps limit mutation accumulation because only a small fraction of all cells (~5%) in the crypt are stem cells. Mutations in non-stem cells normally cannot accumulate because these cells are lost within a week. Within stem cells, a primary defense against mutations is DNA replication fidelity, with the immortal strand hypothesis (3) an extreme scenario where any replication errors are subsequently

lost because newly synthesized DNA strands are asymmetrically distributed to non-stem cell daughters. Another mechanism of mutation avoidance is stem cell quiescence, where stem cells divide infrequently. Stem cells may also be extremely sensitive to DNA damage, and therefore mutant stem cells could be eliminated by apoptosis (4).

Recent mouse studies illustrate that Lgr5+ intestinal stem cells are not quiescent, with mitotic rates estimated at once per day (5). However, there may be many different types of intestinal stem cells, and potentially some of these are more quiescent (6). The immortal strand hypothesis (3) in intestinal crypts is controversial, with evidence for and against asymmetric DNA strand segregation [see for example references (7–12)].

Human crypt stem cell studies are limited because of the inability to use the powerful fate mapping experimental techniques of model systems. However, humans are long-lived and therefore it is possible to measure the “success” of fidelity mechanism by directly measuring mutations in colons isolated from different aged individuals (13). Measurable mutations should accumulate over decades of aging, even with relatively low mutation or stem cell division rates. Human stem cell fidelity mechanisms can

be inferred from the numbers and types of somatic mutations accumulated over a lifetime.

A practical problem with measuring somatic mutations is that because different mutations occur in different cells, the frequency of any specific mutation in a population of human cells is likely to be too low (<5%) to measure with most techniques. Crypts contain multiple stem cells (2, 6) and a mutation in one stem cell would be masked by surrounding non-mutant cells in the same crypt. However, progeny of a single mutant stem cell may expand and fill the crypt because stem cells are extrinsically defined by a surrounding niche, and their survival is probabilistic. Niche stem cells usually divide asymmetrically to produce one stem and one non-stem cell daughter, but may also divide symmetrically to produce two stem cell daughters (expansion) or two non-stem cell daughters (extinction). The total number of niche stem cells remains constant, but eventually all stem cell lineages except one become extinct (**Figure 1**) by a niche succession process called neutral drift (14, 15). Crypt niche stem cell neutral drift has been characterized in mice using fate marking strategies that are impractical in humans. However, measurements of passenger methylation patterns in human crypts are also consistent with multiple stem cells per crypt and neutral drift, with niche succession intervals estimated at about 8 years (16). This human crypt niche succession time estimate is uncertain because of a paucity of experimental opportunities to characterize human crypt niche dynamics. Another study also found that crypt stem cell succession times are longer in humans compared to mice (17).

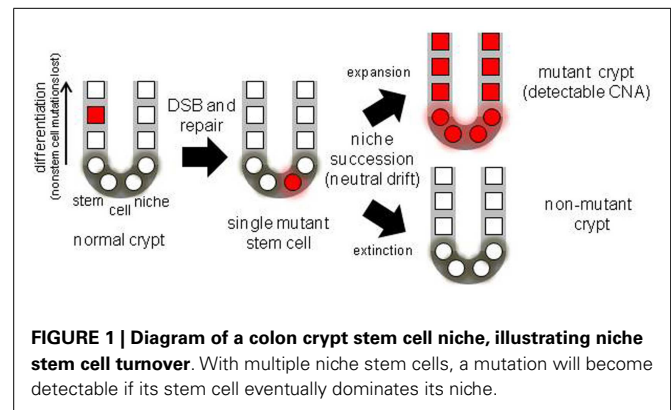
A mutant crypt stem cell can either suffer extinction or expand to fill the entire niche (fixation), and become detectable (**Figure 1**). Neutral drift appears to constantly recur in the absence of mutations (14, 15), but potentially a mutation may influence this process by conferring positive or negative selection to its cells. Therefore, crypt mutation frequencies reflect many aspects of stem cell genetic fidelity, from replication fidelity to the ultimate fates of mutant stem cell progeny. Here data (13) from crypt genomes scanned with high density SNP microarrays for chromosomal copy number alterations (CNAs) from individuals of different ages are further analyzed.

MATERIALS AND METHODS

COLON CRYPTS AND ANALYSIS

The data are from Ref. (13), using the “Reference” method with additional analysis (see below). Briefly, single individual whole normal crypts were obtained from $\sim 1 \text{ cm}^2$ mucosal patches of fresh colectomies at the University of Southern California Keck School of Medicine using an EDTA washout method (16). Procurement of the excess tissue was approved by the Institutional Review Board at the University of Southern California. Normal crypts were isolated from normal appearing colon obtained at least 10 cm away from a tumor. DNA was extracted in 15 μl of TE with 1 μl of 20 mg/ml Proteinase K at 56°C for 4 h followed by boiling for 7 min. All this DNA was used for the SNP microarrays (610-Quad, 660-Quad, 730-OmniExpress) using standard Illumina protocols.

Data were processed using GenomeStudio with a quality threshold of 0.15. Call rates for 180 crypts from 18 colons were variable (49.5–99.8%, average 89.5%), likely because each crypt



has $\sim 20 \text{ ng}$ of DNA, versus the recommended 200 ng of DNA per microarray. Crypts with call rates greater than 60% ($N = 176$) were further analyzed. Multiple crypts from the same colon were compared pairwise at the SNPs. The genotype of a “reference” crypt (typically one with a higher call rate) was determined with GenomeStudio. The reference data were filtered by eliminating no call SNPs and the homozygous (AA, BB) loci. The filtered reference crypt data were compared pairwise with the subject crypt. A likely CNA was identified by loss of heterozygosity (LOH) in the subject crypt at a string of adjacent heterozygous SNPs. At least two different reference crypts were used.

The percentages of SNPs with LOH outside of these likely CNAs were less than 2% of the AB SNPs, even with lower call rates. Most of these non-CNA LOH SNPs were singletons, with fewer longer strings. The probabilities that LOH occurred by chance in the likely CNAs were calculated using the Poisson distribution, with error probabilities estimated from non-CNA LOH SNP frequencies. The smallest CNA had five adjacent heterozygous SNPs and the probabilities that the CNAs < 1 Mb occurred by chance were all less than 0.01 and typically much smaller. All CNAs were verified by manual inspection.

Log R ratios were used to distinguish between deletions or gene conversion (GC). A deletion was called when the average log R ratio of the LOH region was more than 0.25 lower than its flanking regions without LOH. Duplications were not formally analyzed for because of the wide variations in log R ratios, but were visually identified as large regions with three chromosome copies with increased log R ratios.

ADDITIONAL ANALYSIS

The crypt SNP microarray data were further reanalyzed with Nexus Copy Number software (Version 7, BioDiscovery, Hawthorne, CA, USA) using the SNP-FASST2 analysis with default settings and a paired crypt analysis. The software did not identify all CNAs identified in Ref. (13), but all of the previously identified CNAs were evident with manual inspection. One additional deletion on chromosome 1q in an 85-year-old male was identified by the software, which was missed by the previous analysis because the mutant cells were a minority of all cells, with heterozygous SNPs still called AB by GenomeStudio.

The proportions of mutant cells within a crypt containing a deletion were estimated by comparing average BAFs of the subject

crypt that were homozygous (BAF_{AA} or BAF_{BB}) in the reference crypt, versus SNPs that became homozygous (BAF_{aa} or BAF_{bb}). Simplistically, if all cells within a crypt contained the deletion, then BAFs at the previous AB SNPs would be identical to the SNPs that were always AA or BB. The formula used was,

$$P = 1 - \left[\frac{(\text{BAF}_{AA} - \text{BAF}_{aa})}{(0.5 - \text{BAF}_{AA})} + \frac{(\text{BAF}_{BB} - \text{BAF}_{bb})}{(\text{BAF}_{BB} - 0.5)} \right] \cdot \frac{1}{2}.$$

RESULTS

CHROMOSOME COPY NUMBER ALTERATIONS INCREASE WITH AGE

Detectable human crypt CNAs increased with age (Table 1 and Figure 2). No CNAs were found in 48 crypts from individuals less than 50 years old, with 14 CNAs in 13 of 89 crypts (15%) from individuals greater than 50 years of age ($p = 0.0042$, Fisher's exact test). Call rates were not significantly different between crypts with and without CNAs. CNAs sizes (Table 2) ranged from ~10,000 bp (LOH at five adjacent SNPs) to 98 Mb (LOH at 3,996 SNPs). Only one crypt had more than one CNA. Seven CNAs were greater than 1 Mb, and six were smaller than 1 Mb (Figure 2B). A single 44 Mb interstitial duplication was detected. The CNAs appeared to be in the majority (average 93%, range 34–100%) of cells within each crypt (Table 2). Only one LOH region was estimated to be in less than 80% of the crypt cells. Therefore, most of the CNAs were fixed or near fixations in their crypts.

The CNAs were deletions, GC events, and one duplication (Table 2). Log R ratios were consistent with LOH by GC for the three largest CNAs that involved nearly the whole 1q chromosome arm (98 Mb) and the distal ends of 17q (46 Mb) and 7q (24 Mb). These large CNAs included their telomeres, and could be generated by a single double strand DNA break (DSB). The ten other CNAs with reduced log R ratios appeared to be simple deletions generated by two DSBs. The single interstitial duplication would also require at least two DSBs.

DIFFERENT TYPES OF CHROMOSOME COPY NUMBER ALTERATIONS IN INFLAMMATORY BOWEL DISEASE

Normal colon crypts ($N = 36$) were also analyzed from four patients with ulcerative colitis (UC), which is characterized by inflammation and regeneration (18). There were more UC CNAs with 39% of crypts having CNAs (Tables 1 and 2). Five crypts had two or more different CNAs. All UC CNAs appeared to be deletions (decreased log R ratios), and were present in more than 85% of the crypt cells. Call rates were not significantly different between UC crypts with and without CNAs.

The UC CNAs were different from the CNAs in non-UC crypts (Figures 2A,C). A single large deletion (~12.4 Mb) was found in one UC crypt, but the 21 other UC CNAs were small (<1.0 Mb) non-identical deletions clustered at two “hotspots” at 3p14.2 and 16p13.3. Four UC crypts had multiple 3p14.2 or 16p13.3 deletions, some on both the maternal and paternal alleles, resulting in homozygous deletions (HD) in two crypts. The 3p14.2 deletion ($N = 6$) was present in the FHIT locus, a known DNA fragile site (19). The 16p13.3 deletion was more common ($N = 15$), and was also observed twice in the non-UC colon from patient 13. The

Table 1 | Colon crypt chromosome copy number alterations.

Individual	Age/sex	Disease	Crypts	Mutant crypts (%)	Total CNAs
NORMAL					
1	17/M	CRC	6	0 (0)	0
2	26/F	CRC	5	0 (0)	0
3	27/F	Diverticulitis	10	0 (0)	0
4	28/F	CRC	10	0 (0)	0
5	36/F	Endometriosis	10	0 (0)	0
6	45/M	CRC	7	0 (0)	0
7	57/M	CRC	10	2 (20)	2
8	72/M	CRC	11	0 (0)	0
9	78/F	Diverticulitis	9	1 (11)	1
10	80/M	CRC	14	1 (7)	1
11	83/M	CRC	16	0 (0)	0
12	85/M	CRC	12	2 (17)	2
13	89/F	CRC	14	6 (43)	7
14	98/M	CRC	3	1 (33)	1
Total			137	13 (9)	14
	<50 years		48	0 (0)	0
	>50 years		89	13 (15)	14
COLITIS					
15	30/M	UC	12	6 (50)	8
16	56/M	UC	11	7 (64)	12
17	57/F	UC	8	1 (12)	1
18	46/F	UC	7	1 (14)	1
Total			38	15 (39)	22

CRC, colorectal cancer; UC, ulcerative colitis.

16p13.3 deletion is in a region commonly deleted in cancer cell lines (“16p 6 Mb unexplained”), that also appears to be a DNA fragile site (20). Therefore, DSBs at two specific DNA fragile sites are common in normal UC crypts.

EXPECTED CHROMOSOME COPY NUMBER FREQUENCIES WITH AGE

The increases in crypt CNAs with aging may reflect abnormal losses of stem cell fidelity or could represent mutation frequencies expected with normal mutation and cell division rates. It is possible to calculate how CNA frequencies should increase with age if one knows the normal error and division rates of crypt stem cells. The CNA mutation rate in normal cells is uncertain but chromosomal instability (CIN) of ~0.001 chromosomal changes per division have been measured in colorectal cancer (CRC) cell lines (21, 22). Assuming replication fidelity is ~10–100X higher in normal cells (23–25), stem cell division every day to once a week, and a single stem cell per crypt, expected frequencies of mutant crypts with age were plotted (Figure 3). With this limited data, observed mutant crypt frequencies are not markedly different from that expected with these relatively modest combinations of mutation and division rates. The modeled increase in mutant crypts is roughly linear with age, and a lag between the acquisition of a CNA in a single stem cell and its subsequent fixation to detectable levels in its niche (estimated at ~8 years in human crypts (16)) can help account for the relative lack of mutant crypts at earlier ages.

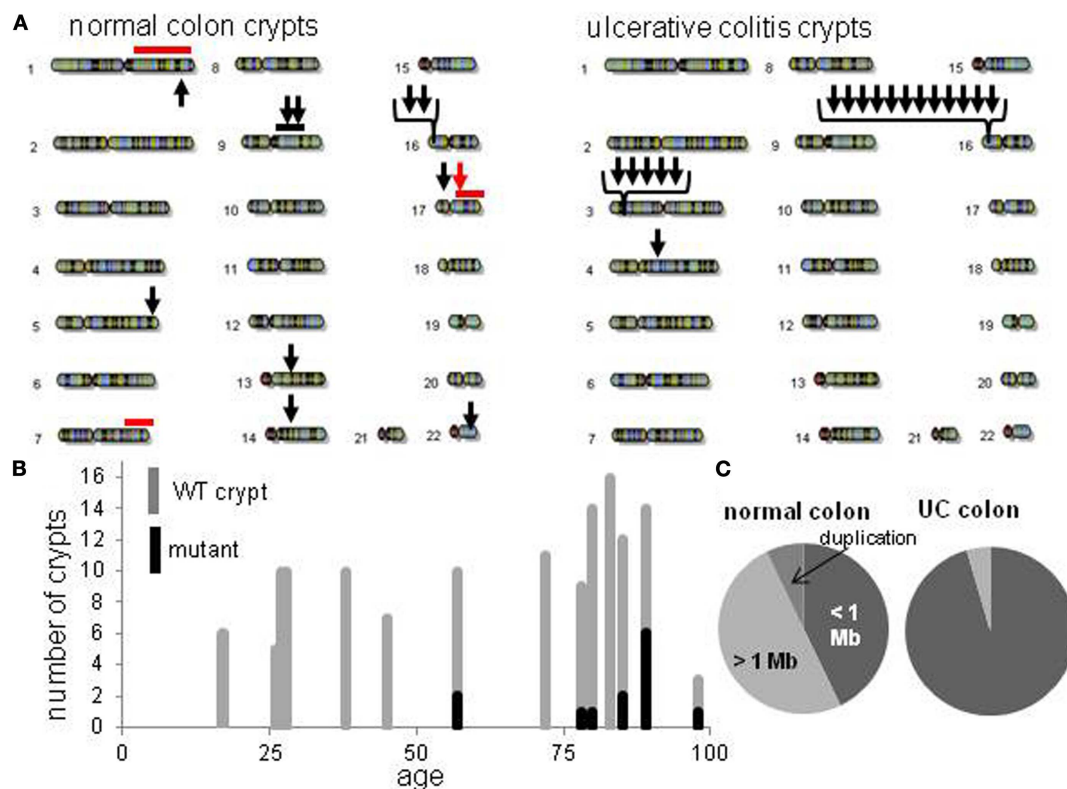


FIGURE 2 | Crypt chromosome copy number alterations (CNAs) across all non-UC and UC crypts. (A) Locations of chromosome CNAs. Arrows indicate loss of heterozygosity, red bars indicate gene conversion, black bar is

a duplication. **(B)** Increase in mutant non-UC crypts with age. **(C)** Relative proportions of chromosome CNA types in normal colon crypts and normal UC crypts.

LACK OF EVIDENCE FOR MUTATION SELECTION IN CRYPT STEM CELL NICHES

The CNAs in normal crypts could reflect selection for mutations that confer selective advantages, or random mutations fixed due to neutral drift. Although it is uncertain which CNAs confer selective advantages to a crypt stem cell, CNAs commonly found in CRCs may predispose to neoplastic progression. We therefore compared the chromosomal locations of the 12 crypt LOH regions (non-fragile sites) to 41 common LOH regions found in 269 MSI negative CRCs from Ref 26 (**Figure 4**). Similar to the non-UC colon crypts, most CRC LOH regions (78%) were greater than 1 Mb. The regions commonly deleted in CRCs include much of the genome (~45% of the genome for $\geq 10\%$ mutation frequencies, and ~33% for $\geq 15\%$ mutation frequencies), and the crypt CNAs may fall within these regions by chance. Of the 12 crypt LOH CNAs, 7 overlapped with CRC LOH CNAs with mutation frequencies $\geq 10\%$, but only 2 overlapped with CRC mutation frequencies $\geq 15\%$ (**Figure 4**). To test whether crypt CNAs were over or under-represented within the CRC CNA regions, a Chi-square test was performed (**Table 3**). The observed crypt CNA locations with respect to the CRC CNAs were not significantly different than expected by chance ($p > 0.05$).

Another study measured CNAs in normal whole human blood cells and also found age-related increases, with ~2–3% detectable

CNA incidence in the elderly (27). Blood cells are mixtures of many different cell types that originate from hematopoietic stem cells in multiple widespread bone marrow niches (28). The detection of a CNA in the blood implies the spread of a mutant hematopoietic stem cell to multiple bone marrow niches. In contrast to the colon crypt CNAs that arise within isolated single small niches, the whole blood CNAs were commonly found within chromosomal regions frequently altered in hematopoietic malignancies and normal individuals with detectable blood CNAs had higher risks for subsequent hematopoietic malignancies.

DISCUSSION

The numbers and types of somatic mutations accumulated over a lifetime reflect different aspects of stem cell fidelity. Mutations are potentially deleterious, and in theory, there may exist special mechanisms that prevent their accumulation within stem cell lineages (3). Primary defenses against somatic mutations are high replication fidelity, efficient DNA repair, and reduced stem cell divisions. The current data indicate that more than 10% of normal human crypts accumulate at least one measurable CNA after the age of 50 years. As illustrated in **Figure 3**, the observed CNA accumulation is consistent with relatively high replication fidelity and low stem cell division rates. Therefore, even with high genetic fidelity, somatic mutations can accumulate because of long human lifetimes. Other studies using histologic markers

Table 2 | Chromosome copy number alteration characteristics.

Crypt	Age	Location ^a	Size (Mb)	Type	Mutant (%)
NORMAL					
7a	57	13q14.11	0.73	Deletion	83
7b	57	17q12	0.015	Deletion	94
9a	78	7q33–q44	24	GC (arm loss)	98
10a	80	1q12–q44	99	GC (arm loss)	96
12a	85	17q12–q25.3	46	GC (arm loss)	99
12b	85	1q43	2.4	Deletion	34
13a	89	5q35.1	3.6	Deletion	95
		9q12–q31.2	44	Duplication	NC
13b	89	9q21.13	0.23	Deletion	89
13c	89	16p13.3	0.073	Deletion	97
13d	89	16p13.3	0.27	Deletion	90
13e	89	17p12	1.1	Deletion	98
13f	89	22q13.1	0.26	Deletion	96
14a	98	9q22.32	3.3	Deletion	99
COLITIS					
15a	30	16p13.3	0.38	Deletion	96
15b	30	16p13.3	0.41	Deletion	95
15c	30	16p13.3	0.34	Deletion	86
15d	30	3p14.2	0.082	Deletion	96
		3p14.2	0.13	Deletion	96
		16p13.3 ^a	0.35	Deletion	89
15e	30	16p13.3 ^a	0.35	Deletion	91
15f	30	3p14.2	0.075	Deletion	98
16a	56	3p14.2	0.14	Deletion	100
16b	56	16p13.3	0.21	Deletion	96
16c	56	16p13.3	0.60	Deletion	97
16d	56	16p13.3	0.20	Deletion (HD)	NC
		16p13.3 ^a	0.15	Deletion (HD)	NC
16e	56	16p13.3	0.010	Deletion	98
		16p13.3	0.11	Deletion	98
16f	56	3p14.2	0.11	Deletion	100
		16p13.3	0.44	Deletion (HD)	NC
		16p13.3 ^a	0.24	Deletion (HD)	NC
16g	56	4q21.1–q22.1	12	Deletion	94
		16p13.3	0.19	Deletion	96
17a	57	16p13.3	0.32	Deletion	98
18a	46	3p14.2	0.042	Deletion	95

^aOther 16p allele within the same colon.

NC, not calculated.

have also demonstrated age-related increases in human crypt mutations (29, 30).

Potentially the observed CNAs may have occurred secondary to losses of normal stem cell fidelity. A cell with increased genetic instability would be expected to accumulate multiple mutations. For example, CRC genomes typically have multiple CNAs (25, 26). However, generally only a single CNA was found in each crypt, which is more consistent with random mutation rather than a crypt specific decrease in genetic fidelity.

LACK OF STEM CELL ENVIRONMENTAL BUFFERING

Stem cell genomes may be protected from environmental stresses. However, the increased CNA frequencies in UC crypts illustrate

that the local microenvironment can influence stem cell genetic fidelity. UC is characterized by severe inflammation with tissue damage and regeneration (18). Potentially increased cell proliferation would simply result in “accelerated” stem cell aging, with more but the same types of CNAs observed in older non-UC crypts. The current data indicate a distinct UC CNA signature characterized by high frequency small (<1 Mb) deletions at two specific DNA fragile sites. Fragile site deletions likely reflect replication stress (31), which is consistent with the higher proliferation of UC. There are multiple known human DNA fragile sites (31) and it is uncertain why only two such sites were commonly altered in normal UC crypts. This distinct mutation signature illustrates that stem cell genomes are sensitive to their microenvironments,

and that UC may increase cancer risks by increasing the numbers of specific mutation types.

STEM CELL ARCHITECTURE: MANAGING MUTATIONS BY PLAYING DICE

Stem cell genetic fidelity is high but appears insufficient to protect against detectable mutation accumulation or chronic environmental stresses, especially over decades. Given that mistakes are inevitable, other downstream mechanisms may minimize the unwanted consequences of either deleterious or beneficial mutations. Such a secondary line of defense is the probabilistic nature of a stem cell niche architecture. Stem cells are a fraction ($\sim 5\%$) of all crypt cells, so most replication errors occur in non-stem cells and are lost. More importantly, there are multiple stem cells per crypt that are extrinsically defined by a niche. These stem cells normally turnover such that eventually progeny of only one current stem cell occupy the entire niche (Figure 1). This stem cell turnover is an important downstream mechanism for managing genetic infidelity. An unwanted consequence of lethal mutations is tissue

loss. A crypt maintained by a single stem cell would be extremely vulnerable to lethal mutations. A crypt maintained by multiple immortal stem cells that always divided asymmetrically would also lack a mechanism to compensate for the death of its stem cells. Multiple niche stem cells protect the crypt against the deleterious effects of lethal mutations because the loss of any stem cell is readily compensated by the expansion (symmetrical division) of a neighboring stem cell lineage.

An important question is how the dominant niche stem cell is chosen. In the absence of mutation, all niche stem cells are similar, and episodic succession occurs through neutral drift (14, 15). With mutation, selection could have both desirable and undesirable consequences. Stem cells with non-lethal deleterious mutations would be eliminated by purifying selection, which would mitigate aging. However, stem cells with beneficial mutations would become dominant, which could predispose to tumorigenesis.

A niche with multiple neighboring stem cells might appear to be an ideal Darwinian setting to discriminate between even minor fitness differences. Selection could impose ratchet-like increases in fitness, but the opposite typically occurs – tissues degenerate with age. How can less fit stem cells dominate their niche? One way to suspend Darwin is through an interesting non-Darwinian phenomenon (32). According to population genetics theory, the role of chance or drift becomes much more important as population size decreases (33, 34). In smaller populations, it becomes increasingly harder to eliminate deleterious mutations or to fix beneficial mutations. Although many parameters influence the balance between chance and selection, generally chance becomes increasingly more important as population sizes slip below one thousand. Crypts stem cell populations are small (<100 stem cells per niche) and therefore chance rather than selection may more determine what types of mutations are fixed (Figure 5). The regions of LOH acquired during human aging did not preferentially fall within regions commonly deleted in CRCs (Figure 4), suggesting random mutation fixation rather than selection.

The numbers of documented somatic CNAs are small, but random mutation fixation due to small niche stem cell numbers may also help explain why neutral passenger mutations are common in carcinomas (35). Many cancers appear later in life, and many alterations found in CRC genomes appear to arise in normal colon

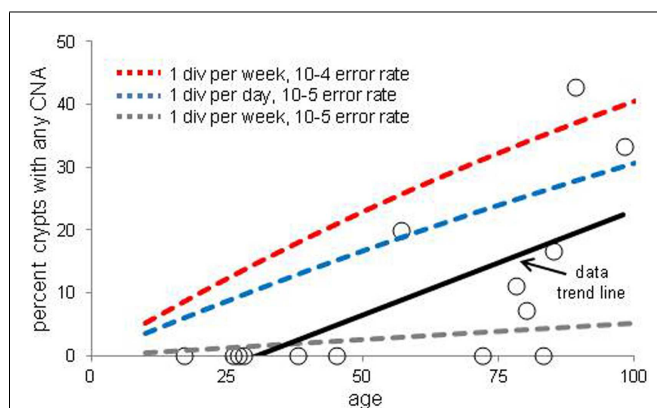


FIGURE 3 | Percent of crypts with any chromosome copy number alteration versus age. Circles are averaged experimental data and the dotted lines indicate calculated mutation frequencies with different combinations of mutation and stem cell division rates. A regression analysis indicates that the experimental increase in CNA frequency with age is significant ($p=0.01$). The calculated mutation frequencies do not account for the lag between a mutation in a stem cell and the time needed to become detectable by niche succession.

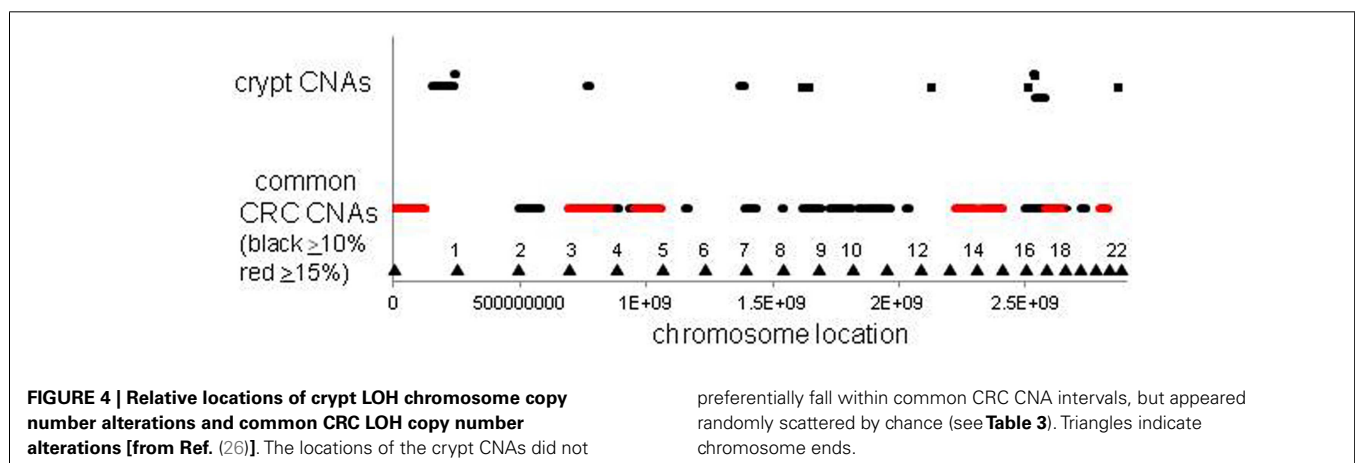


FIGURE 4 | Relative locations of crypt LOH chromosome copy number alterations and common CRC LOH copy number alterations [from Ref. (26)]. The locations of the crypt CNAs did not

preferentially fall within common CRC CNA intervals, but appeared randomly scattered by chance (see Table 3). Triangles indicate chromosome ends.

before visible tumorigenesis (36, 37). Interestingly, cancer genome mutation frequencies are consistent with relatively normal division and mutation rates (38), suggesting many CRC mutations first accumulate in normal crypts. Importantly, there is a profound lack of evidence for purifying selection in many types

of cancer genomes (breast, CRC, pancreatic, glioblastoma, head and neck, ovarian, myeloma, gastric), manifested by the ratio of non-synonymous to synonymous mutations (39, 40). This ratio (dN/dS) is about one and essentially the value expected of random mutation, suggesting that most coding mutations in cancer genomes were not screened by selection. This lack of evidence for somatic mutation selection is curious because the dN/dS ratio is less than one in the human germline (40), indicating that purifying selection normally eliminates many non-synonymous mutations in human populations. The abundance of cancer passenger mutations and the lack of purifying selection may be the legacy of their origins within very small stem cell niches, where mutation selection is nullified by chance fixation.

If non-Darwinian mutation fixation depends on small numbers of niche stem cells, tissues with different niche architectures or dynamics may more often accumulate selective mutations. Whole blood cells originate from multiple hematopoietic stem cell niches that are much more dynamic than crypt niches with respect to physical locations, numbers of stem cells, and migration of stem cells to neighboring niches (28). Instead of a physical subdivision into multiple distinct isolated small stem cell niches, hematopoietic stem cells are not confined to a single niche but normally migrate to new niches. Consistent with a different niche architecture, age-related increases in detectable CNAs in whole blood

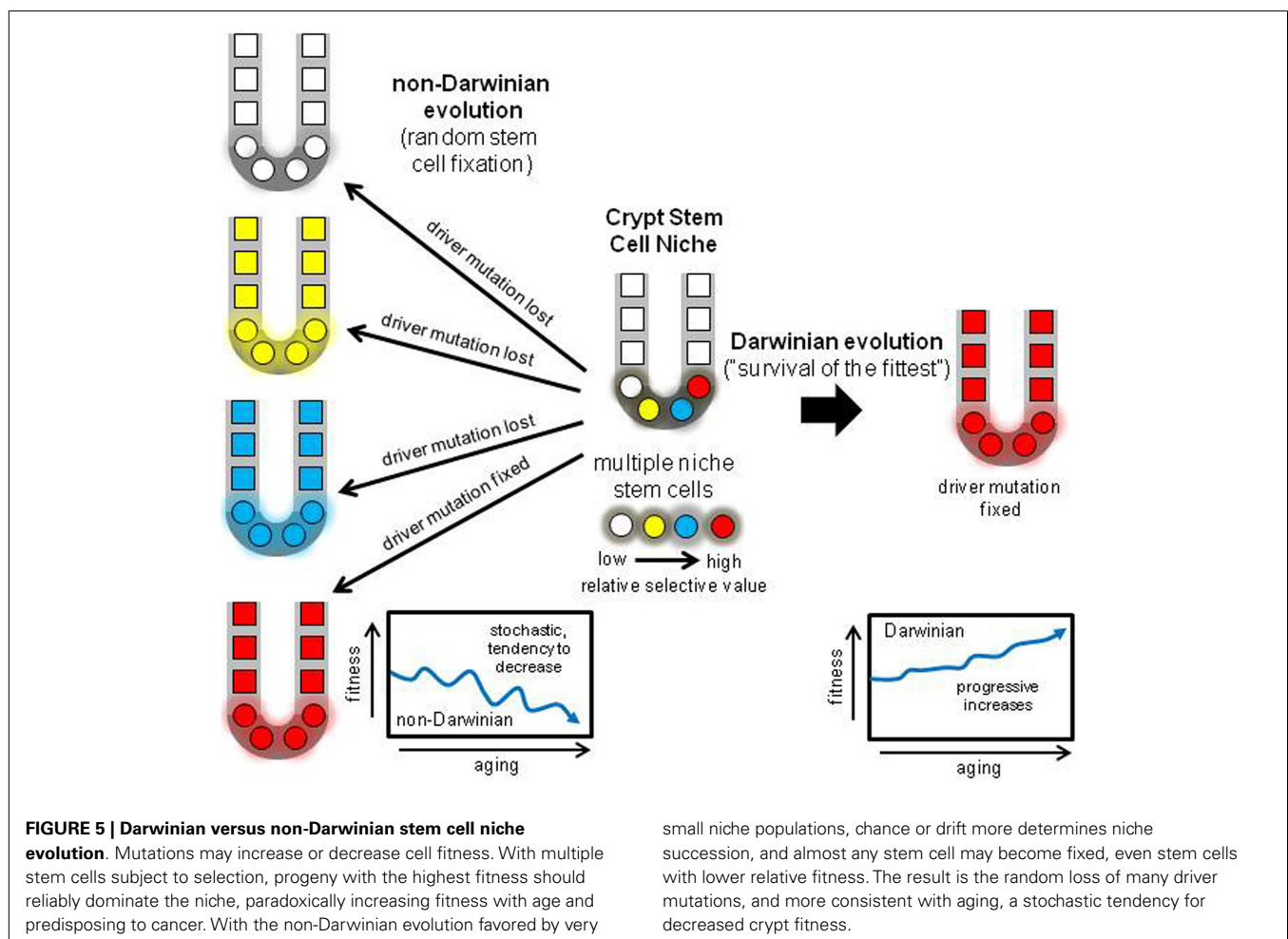
Table 3 | Crypt chromosome copy number alterations versus common colorectal chromosome copy number alterations.

	CRC LOH 10% frequency, $p = 0.353^*$		CRC LOH 15% frequency, $p = 0.231$	
	Observed	Expected ^a	Observed	Expected ^b
Crypt LOH in CRC LOH intervals	7	5.4	2	4
Crypt LOH outside CRC LOH intervals	5	6.6	10	8

*Two tailed p value, Chi-square.

^aCRC LOH regions ($\geq 10\%$ mutation frequencies) cover $\sim 45\%$ of genome.

^bCRC LOH regions ($\geq 15\%$ mutation frequencies) cover about $\sim 33\%$ of genome.



appear to be more driven by selection because blood CNAs are frequently located in regions commonly altered in hematopoietic malignancies (27). Niche stem cell number size limitations may not apply to the hematopoietic system, and therefore selection may have a greater role in determining whether a mutant stem cell can spread and occupy the majority of hematopoietic stem cell niches.

Random niche mutation fixation is not a fool-proof anti-cancer mechanism because by chance sometimes a potential driver mutation will become fixed instead of discarded, and some somatic mutations may confer selection sufficient to overwhelm random fixation. Many driver mutations such as in APC and TP53 are compatible with normal appearing intestines, and their fixation within a crypt resembles clonal evolution, with a net increase in mutant cells but without visible tumorigenesis. Indeed, some “driver” mutations without immediate apparent selective value may be randomly fixed, expressing their driver functions only in combination with other driver mutations much later in the final tumor or metastasis. Potentially CRCs could result from the random accumulation, in any order, of multiple initially neutral, “driver” mutations in niche stem cells (41). However fewer CRCs would occur with multiple crypt niche stem cells compared to crypts with multiple immortal stem cells (42).

NON-DARWINIAN STEM CELL NICHE EVOLUTION: A TESTABLE HYPOTHESIS

The hypothesis that stem cell niches harness non-Darwinian evolution can be tested experimentally in model systems by comparing the fates of specific mutations engineered to occur in single isolated stem cells. Mouse crypt niches are likely smaller than human niches, so non-Darwinian effects should be exacerbated. For example, with a mouse model with a mutant *Cre* sporadically reactivated by rare back-mutation, the fixation of an individual intestinal crypt stem cell with a neutral floxed *lacZ* marker (*Rosa26R*) can be compared to the fixation of a stem cell with *Rosa26R* combined with floxed “driver” mutations (43, 44). With Darwinian evolution, driver mutations should confer selective advantages and be fixed much more often, leading to more *lacZ* positive crypts. With non-Darwinian evolution, stem cells with driver mutations should be randomly discarded as often as stem cells with neutral *Rosa26R* mutations, resulting in similar numbers of *lacZ* positive crypts. Predicted differences between niche selection and random fixation are large. With “*N*” niche stem cells (*N* is about 8–12 stem cells per crypt in mice (14, 15), stem cell fixation should be 100% with driver mutation selection, but only 8–12% (“1/*N*”) of stem cells will become fixed with random stem cell loss.

Data with floxed *Kras*^{G12D} and *Apc*^{580S} driver mutations were more consistent with non-Darwinian evolution or random niche fixation because the numbers of fixed *lacZ* positive mutation events were similar to control mice without the driver mutations (43, 44). Although stem cells with *Kras*^{G12D} or *Apc*^{580S} mutations did not appear to be fixed more often in crypt niches, they did confer selective advantages after fixation, manifested by larger patches of mutant crypts due to increased crypt fission. Similar experimental studies can further test whether isolated single niche stem cells with specific somatic mutations are fixed randomly or selectively.

CRYPT STEM CELL GENETIC FIDELITY AND NON-DARWINIAN STEM CELL EVOLUTION

Perfect stem cell fidelity would be an “anti-evolution” strategy to never grow old. Aging, or the accumulation of mutations may be inevitable, and the genetic fidelity of human crypt stem cells appears not to be higher than expected of normal cells. Given the inevitability of mutations, the crypt stem cell niche may trade Darwinian for non-Darwinian evolution as a downstream mechanism to manage these mutations (Figure 5). During a lifetime, a critical question is whether deleterious or beneficial mutations are more dangerous to homeostasis. Many “deleterious” somatic mutations may be tolerated by human cells, exemplified by the relatively large numbers of rare but potentially dysfunction mutations in normal human germline genomes (45). The spread of beneficial somatic mutations may pose a greater threat to survival. Niche stem cell turnover may harness a non-Darwinian evolution mechanism (neutral drift) that readily protects against lethal mutations and helps ensure that beneficial mutations that might lead to cancer are often discarded. Given the cooperation needed between multiple cells in mammalian tissues and the dangers of tumorigenesis, an optimal reliable downstream strategy to guard against the unwanted effects of some mutations may be random stem cell fixation in tissues subdivided into very small niches. This non-Darwinian strategy is built into the tissue niche architecture from birth, and can help explain why tissues do not become paradoxically “fitter” with age (Figure 5). This scenario resembles Muller’s ratchet, where asexual division leads to decline (46).

Interesting, non-intuitive phenomenon often emerge at smaller physical dimensions. Multiple, mitotic stem cells in very small niches with non-Darwinian evolution can better explain colon aging and somatic mutation frequencies and spectra compared to Darwinian niche selection. Non-Darwinian evolution may predominate whenever reproducing somatic tissues are physically subdivided into distinct very small isolated compartments. Given the impracticality of human experimental manipulations, the analysis of somatic alterations found in normal human tissues provides a feasible pathway for insights into human stem cell mechanisms. Somatic alterations can reveal much about stem cell life and death, particularly because most mitotic niche stem cell lineages suffer extinction. Newer technologies increasingly provide better methods to detect mutations, and more data on the numbers and types of alterations found in normal human tissues will allow much better inferences on how we age. Small stem cell niches provide a downstream architectural mechanism for randomly discarding many inevitable but unwanted mutations.

AUTHORS CONTRIBUTION

Haeyoun Kang helped analyze the data and edit the manuscript. Darryl Shibata supplied the colon crypts, helped analyze the data, and write the manuscript.

ACKNOWLEDGMENTS

The authors acknowledge the assistance of authors in Ref. (13) (John C. F. Hsieh, Chih-Lin Hsieh, David Van Den Berg, and Michael R. Lieber) who generously allowed access to their primary data, and the technical assistance of Renae K. Shibata.

REFERENCES

- Potten CS, Loeffler M. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* (1990) **110**:1001–20.
- Humphries A, Wright NA. Colonic crypt organization and tumorigenesis. *Nat Rev Cancer* (2008) **8**:415–24. doi:10.1038/nrc2392
- Cairns J. Mutation selection and the natural history of cancer. *Nature* (1975) **255**:197–200. doi:10.1038/255197a0
- Potten CS. Extreme sensitivity of some intestinal crypt cells to X and gamma irradiation. *Nature* (1977) **269**:518–21. doi:10.1038/269518a0
- Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, Cozijnsen M, et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* (2007) **449**:1003–7. doi:10.1038/nature06196
- Clevers H. The intestinal crypt, a prototype stem cell compartment. *Cell* (2013) **154**:274–84. doi:10.1016/j.cell.2013.07.004
- Potten CS, Hume WJ, Reid P, Cairns J. The segregation of DNA in epithelial stem cells. *Cell* (1978) **15**:899–906. doi:10.1016/0092-8674(78)90274-X
- Potten CS, Owen G, Booth D. Intestinal stem cells protect their genome by selective segregation of template DNA strands. *J Cell Sci* (2002) **115**:2381–8.
- Quyn AJ, Appleton PL, Carey FA, Steele RJ, Barker N, Clevers H, et al. Spindle orientation bias in gut epithelial stem cell compartments is lost in precancerous tissue. *Cell Stem Cell* (2010) **6**:175–81. doi:10.1016/j.stem.2009.12.007
- Schepers AG, Vries R, van den Born M, van de Wetering M, Clevers H. Lgr5 intestinal stem cells have high telomerase activity and randomly segregate their chromosomes. *EMBO J* (2011) **30**:1104–9. doi:10.1038/emboj.2011.26
- Escobar M, Nicolas P, Sangar F, Laurent-Chabalier S, Clair P, Joubert D, et al. Intestinal epithelial stem cells do not protect their genome by asymmetric chromosome segregation. *Nat Commun* (2011) **2**:258. doi:10.1038/ncomms1260
- Steinhauser ML, Bailey AP, Senyo SE, Guillermier C, Perlstein TS, Gould AP, et al. Multi-isotope imaging mass spectrometry quantifies stem cell division and metabolism. *Nature* (2012) **481**:516–9. doi:10.1038/nature10734
- Hsieh JC, Van Den Berg D, Kang H, Hsieh CL, Lieber MR. Large chromosome deletions, duplications, and gene conversion events accumulate with age in normal human colon crypts. *Aging Cell* (2013) **12**:269–79. doi:10.1111/ace.12053
- Lopez-Garcia C, Klein AM, Simons BD, Winton DJ. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* (2010) **330**:822–5. doi:10.1126/science.1196236
- Snippert HJ, van der Flier LG, Sato T, van Es JH, van den Born M, Kroon-Veenboer C, et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* (2010) **143**:134–44. doi:10.1016/j.cell.2010.09.016
- Yatabe Y, Tavaré S, Shibata D. Investigating stem cells in human colon by using methylation patterns. *Proc Natl Acad Sci U S A* (2001) **98**:10839–44. doi:10.1073/pnas.191225998
- Campbell F, Williams GT, Appleton MA, Dixon MF, Harris M, Williams ED. Post-irradiation somatic mutation and clonal stabilisation time in the human colon. *Gut* (1996) **39**:569–73. doi:10.1136/gut.39.4.569
- O'Connor PM, Lapointe TK, Beck PL, Buret AG. Mechanisms by which inflammation may increase intestinal cancer risk in inflammatory bowel disease. *Inflamm Bowel Dis* (2010) **16**:1411–20. doi:10.1002/ibd.21217
- Saldivar JC, Shibata H, Huebner K. Pathology and biology associated with the fragile FHIT gene and gene product. *J Cell Biochem* (2010) **109**:858–65.
- Bignell GR, Greenman CD, Davies H, Butler AP, Edkins S, Andrews JM, et al. Signatures of mutation and selection in the cancer genome. *Nature* (2010) **463**:893–8. doi:10.1038/nature08768
- Ribas M, Masramon L, Aiza G, Capellà G, Miró R, Peinado MA. The structural nature of chromosomal instability in colon cancer cells. *FASEB J* (2003) **17**:289–91.
- Roschke AV, Stover K, Tonon G, Schäffer AA, Kirsch IR. Stable karyotypes in epithelial cancer cell lines despite high rates of ongoing structural and numerical chromosomal instability. *Neoplasia* (2002) **4**:19–31. doi:10.1038/sj.neo.7900197
- Thompson SL, Compton DA. Examining the link between chromosomal instability and aneuploidy in human cells. *J Cell Biol* (2008) **180**:665–72. doi:10.1083/jcb.200712029
- Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature* (1998) **396**:643–9. doi:10.1038/25292
- Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancers. *Nature* (1997) **386**:623–7. doi:10.1038/386623a0
- Xie T, D'Ario G, Lamb JR, Martin E, Wang K, Tejpar S, et al. A comprehensive characterization of genome-wide copy number aberrations in colorectal cancer reveals novel oncogenes and patterns of alterations. *PLoS ONE* (2012) **7**:e42001. doi:10.1371/journal.pone.0042001
- Laurie CC, Laurie CA, Rice K, Doherty KF, Zelnick LR, McHugh CP, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet* (2012) **44**:642–50. doi:10.1038/ng.2271
- Morrison SJ, Spradling AC. Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell* (2008) **132**:598–611. doi:10.1016/j.cell.2008.01.038
- Campbell F, Fuller CE, Williams GT, Williams ED. Human colonic stem cell mutation frequency with and without irradiation. *J Pathol* (1994) **174**:175–82. doi:10.1002/path.1711740306
- Greaves LC, Preston SL, Tadrous PJ, Taylor RW, Barron MJ, Oukrif D, et al. Mitochondrial DNA mutations are established in human colonic stem cells, and mutated clones expand by crypt fission. *Proc Natl Acad Sci U S A* (2006) **103**:714–9. doi:10.1073/pnas.0505903103
- Durkin SG, Glover TW. Chromosome fragile sites. *Annu Rev Genet* (2007) **41**:169–92. doi:10.1146/annurev.genet.41.042007.165900
- King JL, Jukes TH. Non-Darwinian evolution. *Science* (1969) **16**:788–98. doi:10.1126/science.164.3881.788
- Whitlock MC, Bürger R. Fixation of new mutations in small populations. In: Ferrière R, Dieckmann U, Couvet D editors. *Evolutionary Conservation Biology*. Cambridge: Cambridge University Press (2004). p. 155–70.
- Whitlock MC. Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles, and sexual selection. *Evolution* (2000) **54**:1855–61. doi:10.1111/j.0014-3820.2000.tb01232.x
- Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. *Science* (2013) **339**:1546–58. doi:10.1126/science.1235122
- Tomasetti C, Vogelstein B, Parmigiani G. Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation. *Proc Natl Acad Sci U S A* (2013) **110**:1999–2004. doi:10.1073/pnas.1221068110
- Woo YJ, Siegmund KD, Tavaré S, Shibata D. Older individuals appear to acquire mitotically older colorectal cancers. *J Pathol* (2009) **217**:483–8. doi:10.1002/path.2506
- Wang TL, Rago C, Silliman N, Ptak J, Markowitz S, Willson JK, et al. Prevalence of somatic alterations in the colorectal cancer cell genome. *Proc Natl Acad Sci U S A* (2002) **99**:3076–80. doi:10.1073/pnas.261714699
- Rubin AF, Green P. Mutation patterns in cancer genomes. *Proc Natl Acad Sci U S A* (2009) **106**:21766–70. doi:10.1073/pnas.0912499106
- Woo YH, Li WH. DNA replication timing and selection shape the landscape of nucleotide variation in cancer genomes. *Nat Commun* (2012) **3**:1004. doi:10.1038/ncomms1982
- Calabrese P, Shibata D. A simple algebraic cancer equation: calculating how cancers may arise with normal mutation rates. *BMC Cancer* (2010) **10**:3. doi:10.1186/1471-2407-10-3
- Calabrese P, Tavaré S, Shibata D. Pre-tumor progression: clonal evolution of human stem cell populations. *Am J Pathol* (2004) **164**:1337–46. doi:10.1016/S0002-9440(10)63220-8
- Miller AJ, Dudley SD, Tsao JL, Shibata D, Liskay RM. Tractable Cre-lox system for stochastic alteration of genes in mice. *Nat Methods* (2008) **5**:227–9. doi:10.1038/nmeth.1183
- Fischer JM, Miller AJ, Shibata D, Liskay RM. Different phenotypic consequences of simultaneous versus stepwise Apc loss. *Oncogene* (2012) **31**:2028–38. doi:10.1038/onc.2011.385
- Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, et al. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* (2012) **337**:64–9. doi:10.1126/science.1219240

46. Felsenstein J. The evolutionary advantage of recombination. *Genetics* (1974) **78**:737–56.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that

could be construed as a potential conflict of interest.

Received: 15 August 2013; accepted: 25 September 2013; published online: 14 October 2013.

Citation: Kang H and Shibata D (2013) Direct measurements of human colon crypt stem cell niche genetic fidelity: the

role of chance in non-Darwinian mutation selection. *Front. Oncol.* **3**:264. doi: 10.3389/fonc.2013.00264

This article was submitted to *Cancer Genetics*, a section of the journal *Frontiers in Oncology*.

Copyright © 2013 Kang and Shibata. This is an open-access article distributed under the terms of the Creative

Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



An APC:WNT counter-current-like mechanism regulates cell division along the human colonic crypt axis: a mechanism that explains how *APC* mutations induce proliferative abnormalities that drive colon cancer development

Bruce M. Boman^{1,2*} and Jeremy Z. Fields³

¹ Center for Translational Cancer Research, Helen F. Graham Cancer Center and Research Institute, University of Delaware, Newark, DE, USA

² Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA, USA

³ CATX, Inc., Gladwyne, PA, USA

Edited by:

James L. Sherley, Boston Biomedical Research Institute, USA

Reviewed by:

Nick A. Wright, Barts and the London School of Medicine and Dentistry, UK
Jason C. Mills, Washington University School of Medicine, USA

*Correspondence:

Bruce M. Boman, Center for Translational Cancer Research, Helen F. Graham Cancer Center and Research Institute, University of Delaware, 4701 Ogletown-Stanton Road, Newark, DE 19713, USA
e-mail: brboman@udel.edu

APC normally down-regulates WNT signaling in human colon, and *APC* mutations cause proliferative abnormalities in premalignant crypts leading to colon cancer, but the mechanisms are unclear at the level of spatial and functional organization of the crypt. Accordingly, we postulated a counter-current-like mechanism based on gradients of factors (APC;WNT) that regulate colonocyte proliferation along the crypt axis. During crypt renewal, stem cells (SCs) at the crypt bottom generate non-SC daughter cells that proliferate and differentiate while migrating upwards. The APC concentration is low at the crypt bottom and high at the top (where differentiated cells reside). WNT signaling, in contrast, is high at the bottom (where SCs reside) and low at the top. Given that WNT and APC gradients are counter to one another, we hypothesized that a counter-current-like mechanism exists. Since both APC and WNT signaling components (e.g., survivin) are required for mitosis, this mechanism establishes a zone in the lower crypt where conditions are optimal for maximal cell division and mitosis orientation (symmetric versus asymmetric). *APC* haploinsufficiency diminishes the APC gradient, shifts the proliferative zone upwards, and increases symmetric division, which causes SC overpopulation. In homozygote mutant crypts, these changes are exacerbated. Thus, *APC*-mutation-induced changes in the counter-current-like mechanism cause expansion of proliferative populations (SCs, rapidly proliferating cells) during tumorigenesis. We propose this mechanism also drives crypt fission, functions in the crypt cycle, and underlies adenoma development. Novel chemoprevention approaches designed to normalize the two gradients and readjust the proliferative zone downwards, might thwart progression of these premalignant changes.

Keywords: adenomatous polyposis coli, WNT signaling, survivin, colon cancer stem cells, crypt fission, crypt cycle, adenoma morphogenesis, colonic stem cells

INTRODUCTION

It is well-known that APC down-regulates WNT signaling in normal human colon and that *APC* mutation impairs this down-regulation and contributes to the development of premalignant crypts, which leads to colon cancer [reviewed in (1, 2)]. However, the mechanisms are not well understood at the level of the spatial and functional organization of the colonic crypt. Therefore, we created a counter-current-like model that considers gradients of factors (APC; WNT) along the crypt axis that spatially and temporally regulate colonocyte proliferation and differentiation along this axis. To understand this problem and our proposed solution requires an understanding of the normal colonic crypt.

To better understand the role of APC, crypt renewal, and colonic stem cells (SCs) in maintaining normal form and function of the colon, we will first discuss the organization and function of normal colonic epithelium. This discussion is important because colonic SCs bequeath molecular information to their non-SC progeny that determines the structure and function of normal

colonic epithelium. With that as a foundation, we can then begin to see how changes in populations of SCs can contribute, during colon tumor development, to altered tissue structure and altered tissue function. Although there has been much research on the structure and the function of rodent small intestine, which has increased our understanding of the biology of GI SCs, here we will emphasize knowledge obtained from human colonic SCs, human colonic epithelium, and human colonic cancers. If the reader wishes information in this field as it pertains to SCs in rodent tumorigenesis, several excellent reviews are available (3–5).

HISTOLOGIC AND PROLIFERATIVE CHARACTERISTICS OF NORMAL HUMAN CRYPTS THAT CONTAIN WILD-TYPE *APC*

Anatomically, colonic epithelium in humans is made of regular, pit-like structures called crypts, each containing two to three thousand cells (6, 7). The epithelium of the colon has very high turnover – it is replaced every 5 days through crypt renewal (6). Because the human colon contains $\sim 10^{11}$ cells (8) nearly 10 trillion

colonocytes are generated per year. Remarkably, it is the colonic SC that underlies the generation of this large number of cells during an individual's lifetime while, at any given time, maintains the number of crypt cells constant and crypt dynamics at steady state.

THE ROLE OF NORMAL SCs IN NORMAL HUMAN CRYPT STRUCTURE AND FUNCTION

The bottom of the crypt is where most colonic SCs reside. They are generally quiescent but do generate rapidly proliferating cells (transit-amplifying cells) that are simultaneously differentiating and proliferating (i.e., they are maturing) as they migrate upwards along the crypt axis. As they migrate, they are maturing along the various cell lineages such as absorptive (columnar) goblet (mucin-producing), and other cell types (9–12). Maturing cells, in turn, generate fully mature cells. These terminally differentiated cells continue migrating upwards, become apoptotic, and are eventually sloughed off, at the crypt top, into the colonic lumen.

Several mechanisms are in place to maintain the crypt size and the colonocyte population size constant. In dividing, the SC population at the bottom of the crypt undergoes self-renewal and, at the same time, generates the population of transit-amplifying cells (13). Because colonic SCs are long-lived, they are essential for crypt self-renewal over the lifespan of each individual. Extrapolations of findings from biologic studies in rodents suggests that SCs in a human colonic crypt are a small proportion (~1%) of all cells in that crypt (14). This estimate is in accord with recent immunostaining experiments in human colonic crypts for the SC markers (15, 16). Nevertheless, these rare SCs drive crypt renewal and are key to crypt homeostasis and viability (17, 18).

DYNAMICS OF NORMAL CRYPT CELL POPULATIONS

The dynamics of human colonic crypts are complex [reviewed in (19)]. (i) Crypts contain many cell types. (ii) Most crypt cells have neither a static location nor a static phenotype. As most crypt cell types migrate toward the crypt top, they proliferate and differentiate simultaneously (i.e., they undergo maturation). Eventually they become fully mature, no longer proliferate, become terminally differentiated, and, after apoptosis, are extruded at the crypt top. (iii) Not surprisingly, given the above, cell phenotypes change as colonocytes migrate and mature upwards along the crypt axis and various phenotypic markers show gradient-like distributions.

A few cells of the colonic crypt, the SCs, are different. (i) They don't migrate upwards, remaining, instead, near the crypt bottom. (ii) They are multipotent. Human and rodent studies show that colonic SCs generate several lineages (endocrine cells, absorptive cells, goblet cells). Via tissue renewal, SCs replenish not only their own population, but also, all crypt cell types. (iii) SCs are extremely long-lived. Since crypts are closed systems, crypt cells must be generated by SCs that are already residing in the crypt. Therefore, both the number of cells in the normal crypt and the division of SCs require strict physiological regulation.

STUDYING THE GENERATION OF RAPIDLY PROLIFERATING NON-SCs BY SCs

Because of numerous obstacles to the study of SCs in humans, initial studies focused on the functional properties of these cells. One of the earliest ways used to study SCs and to determine

their anatomic location was pulse-labeling of DNA of rapidly proliferating cells – daughter cells that are produced by SCs. Uptake of bromodeoxyuridine (BrdU) or [^3H]thymidine by these daughter cells in human colonic crypts results in *in vivo* labeling of DNA-synthesizing S-phase cells (6, 20–22). When the fraction (proportion) of S-phase (labeled) cells is plotted against cell position (i.e., against cell level) along the crypt axis, from the crypt bottom to the crypt top, the result is a skewed bell-shaped curve termed the labeling index or LI. In normal colonic crypts, the curve for the LI is low at the crypt bottom (level 1) and top (~ level 82) and maximizes at approximately level 15. Sequential LI profiles were used to track these labeled colonocytes, which showed that they migrate from bottom to top, where they are then extruded. These tracking results indicate that SCs must reside at the crypt bottom. These profiles also indicate that there is a small fraction of cells in S-phase at the bottommost crypt levels (6, 23), where SCs are located. This is also consistent with literature reporting that SCs are relatively quiescent (24–26).

IDENTIFICATION, DISTRIBUTION, AND MODE OF CELL DIVISION OF HUMAN COLONIC SCs

To study important questions such as: what regulates the distribution of SC in the human colonic crypt or what is their type of cell division, it has been necessary to find accurate markers for human colonic SCs. This effort has relied on showing that SC markers fulfill certain criteria – ones that differ somewhat from criteria for establishing SC markers in rodents because validating SC markers by lineage tracing cannot readily be done for human tissues for ethical reasons. Thus, validation in humans generally relies on demonstrating characteristics of self-renewal, tumor-initiating ability, long-term repopulating capability, and capacity for multi-lineage differentiation (27). Based on these criteria several reliable markers (e.g., CD44, CD133, CD166, Musashi 1) have been established for normal and malignant human colonic SCs (15, 28–31).

Our own work (16) led to the discovery that ALDH is a marker for human colonic SCs. We found that ALDH positive colonic cells exhibit the known SC properties of anatomic localization and tumor-initiating ability: (a) immunohistochemistry identified a small subpopulation of ALDH1+ cells (~5%) localized to the bottom of normal crypts (where SC reside) and (b) the Aldefluor assay was used to isolate a subpopulation of malignant colonic cells that generates xenograft tumors (also showing the ability for self-renewal). As few as 25 ALDH+ cells generated tumors while as many as 10,000 ALDH– cells from colorectal cancers (CRCs) did not form xenograft tumors. It was also shown that ALDH+ cells possess the SC features of long-term repopulating ability and multi-lineage differentiation (16). This was done by showing that isolated ALDH+ cells have the ability: (1) to be serially passaged long-term as xenografts in mice (and in colonosphere cultures) with continued isolation of ALDH+ cells, and (2) to differentiate into all of the different cell lineages found in colonic tissues based on histologic evaluation of the tumor xenografts. Similar findings have been published by others (32–34). Taken together, this information is consistent with the conclusion that ALDH1 is a SC marker in the colon of humans.

We then determined staining indices to quantify the distribution of colonic SCs (16). Indices for ALDH+ cells and those marked by other SC markers (CD133, CD44) showed that there is a gradient in the number of SCs upwards along the crypt axis. This gradient is similar to the exponential decrease in stemness with distance from the crypt bottom that we previously reported (23). The reason that this gradient exists is explainable in two ways. (i) In the first explanation, there are, upward along the crypt axis, decreases in the fraction of cells that are SCs. In this view, being a SC is *all-or-none*, and SCs can divide asymmetrically or symmetrically. The division of SCs can occur asymmetrically to produce one SC and one non-SC, symmetrically to produce two identical SCs, or symmetrically to produce two non-SC (35, 36). In theory, division of SCs must, on average, be asymmetric in order to maintain the SC number constant (37). That is, they must produce an average of one SC and one non-SC over all crypt cell divisions. Otherwise, the crypt cell population size will change, as it does in colon tumorigenesis, where SC overpopulation occurs (16).

(ii) In the second explanation, gradual decreases occur, upwards along the crypt axis, in the degree of stemness of each maturing cell. In fact, a radiobiology study (17) suggested that stemness is not all-or-none; rather, stemness is lost gradually. Hence it was postulated (38) that in early generations SCs gradually lose the capacity to function as SCs, and, eventually, all SC potential is lost. Other findings (39, 40) indicate that rodent intestinal SCs generate progenitor cells that have some SC-like properties and that become committed to differentiating along a particular cell lineage. This concept, that there are “intermediate degrees of stemness” is consonant with many of the latest rodent models of small intestinal SC with active SCs and ones that are recruitable, and where the “probability of stemness” represents a gradual change rather than a binary change (24, 41–43). That these progenitor cells seem to have intermediate degrees of stemness supports the idea that cells undergo gradual decreases in their degree of stemness as they migrate along the crypt axis and mature (23).

PROLIFERATIVE CHANGES IN APC MUTANT CRYPTS DURING HUMAN COLON TUMORIGENESIS

The mechanisms by which APC mutations lead to CRC initiation have not been fully elucidated. For instance, it is not clear how a germline APC mutation can initiate intestinal tumors as it clearly does in *Apc^{Min/+}* mice and familial adenomatous polyposis (FAP) patients. Most sporadic CRC cases as well are initiated by APC mutations – such mutations are observed in ~80% of cases of sporadic CRC (44, 45). Inactivation of the second APC allele happens during intestinal adenoma and carcinoma development in both *Apc^{Min}* mice and FAP patients (46). This “second hit” typically results in the total absence of wt-APC protein (47, 48). However, in the case of homozygous mutant APC, the truncated APC protein usually contains some residual functions (discussed below) (48, 49).

Histopathologic studies on APC mutant tissues from FAP patients have been done to investigate how APC mutations might lead to development of CRC. An early finding was that proliferative mechanisms in the colonic crypt become dysregulated. The proliferative alterations were first shown several decades ago using pulse-labeling with BrdU or [³H]thymidine and plotting LI

(labeling index) curves (20–22, 50). For normal-appearing FAP crypts, LI curves are shifted toward the crypt middle, maximizing at about level 20 (in normal colon, the maximum is at level 15) (50). This proliferative shift in normal-appearing (not yet dysplastic) FAP crypts is the earliest-known tissue alteration resulting from a germline mutation in the APC gene. Notably, crypts (e.g., FAP crypts) exhibiting this proliferative abnormality don’t show any microscopically visible changes in histology. Crypts begin to show abnormalities in histology only when they become dysplastic, i.e., during the formation, later, of premalignant adenomas, which have a second hit at the APC locus. For adenomatous crypts from FAP patients, LI curves are shifted even further up the crypt, toward the top (51, 52).

It is possible that the observed shift in the distribution of labeled cells reported in these studies might, in theory, have been caused as a result of variation in the length of the crypt (53–55). However, more comprehensive studies on humans have substantiated that this is not the case. In these studies (56, 57), fresh colonic biopsies from unaffected controls and FAP patients were pulse-labeled *ex vivo* with [³H]Thymidine. Moreover, the distribution of labeled cells was not determined based on “crypt level” (which could vary with crypt length) but on the “proportions” of cells along the crypt axis from bottom to top (which would not vary with crypt length). Zhang et al. (58) used a different approach to map the distribution of proliferating cells – namely using quantitative immunohistochemical (IHC) mapping of Ki67-labeled cells and by plots of staining indices. These mapping results showed that in FAP crypts the population of Ki67+ cells extended upward into the crypt middle as compared to distribution of Ki67+ cells in normal crypts where Ki67+ cells were restricted to the bottom-third. Similar results were found by Mills et al. (59). In adenomas from FAP patients, the shift was even more pronounced; cells staining for Ki67 were mostly found at the top of the crypt or on the luminal surface of the adenomatous epithelium. Thus, results from three independent approaches, quantitative IHC crypt mapping (58, 59), pulse-labeling of crypts *in vivo* (20–22, 50–52), and pulse-labeling of crypts *ex vivo* (56, 57), all provided support for the existence of an upward shift of the proliferative zone in normal-appearing and adenomatous crypts in FAP patients.

Other studies have investigated mitotic cells (rather than [³H]thymidine-labeled or BrdU-labeled cells) in FAP colonic crypts (60, 61). But mitotic indices have not been reported probably because scoring mitoses is difficult due to the small number (<0.5%) of mitotic figures per crypt. Wasan et al. (60) did report, using crypt microdissection, on the highest crypt level at which a mitotic figure was observed and found a modestly higher level in FAP patients than in normal patients. However, the difference was not significant, possibly because he was studying only a small series of FAP patients (*n* = 15). Mills et al. (61) studied a larger series (*n* = 29) using the same technique, crypt microdissection. They found a marked and significant (*p* < 0.0001) increase in the number of mitoses per crypt in FAP crypts (14.2) vs. control crypts (5.6). More recently, we (58) used quantitative IHC mapping of phospho-H3+ cells to measure the distribution of mitotic cells. This approach showed that in FAP crypts the population of mitotic cells extended upward into the crypt middle as compared to normal crypts in which phospho-H3+ cells were located in the

bottom-third. In adenomas from FAP patients, the staining index for phospho-H3+ cells revealed that the shift of mitotic cells was even more pronounced for adenomatous crypts. Thus, these data on mitotic cells are consistent with the proliferative shift (based on LI) observed in FAP.

It was unclear, however, how *APC* mutations generate these earliest-known tissue events during colonic neoplasia development, that is, an upward shift along the crypt axis of the proliferative zone (as indicated by shifts in the LI curve). To investigate mechanisms as to how the proliferative abnormality occurs, we turned to mathematical modeling. Our modeling results (37, 62, 63) clearly demonstrated that only increases in crypt SC number, not alterations in apoptosis, differentiation or cell cycle proliferation of non-SC populations, could accurately simulate the LI shift in FAP crypts. This led us to postulate that the missing link between an *APC* mutation and the LI shift in the initiation of CRC in FAP trait carriers is the overpopulation of crypt SCs.

Biological studies (63, 64) that we did to follow up on our modeling study (37, 62) provided data to support this SC overpopulation mechanism, as have other studies. For example, using methylation pattern diversity, Kim et al. (65) found enhanced SC survival in FAP, which is consistent with SC overpopulation in CRC development. In another study (66), the orphan G protein-coupled receptor GPR49 (LGR5) was found to be overexpressed in primary human colon tumors and LGR5 was then found, in rodent studies (67), to be a SC marker. These research findings led us to determine that ALDH1 is a marker for human colonic SC and allowed us to demonstrate that SC overpopulation occurs due to an *APC* mutation during CRC development (16). Using ALDH1 also allowed us to track SC overpopulation in *APC* mutant tissues during the stepwise progression to CRC development in FAP patient tissues.

Molecular studies have also been done to elucidate APC-based mechanisms that contribute to CRC development [reviewed in (1, 2)]. In normal tissues, the APC protein controls WNT signaling by binding to the β -catenin protein in the cytoplasm, which in turn leads to β -catenin degradation. If *APC* is deleted or mutant, the degradation rate of cytoplasmic β -catenin is diminished. In a small proportion of CRC cases, ones that lack *APC* mutations, β -catenin (*CTNNB1*) mutations are found. Both *CTNNB1* and *APC* mutations activate Tcf4-mediated transcription. The increased levels of cytoplasmic β -catenin lead to increased binding to and activation of Tcf4 (Tcf/Lef) transcription factors, factors that regulate target protein expression and, in turn, cell proliferation and differentiation. For example, Korinek et al. (68) showed, in rodents lacking Tcf4, that epithelial SC compartments become depleted in small intestinal crypts. Other studies (69) showed that Apc modulates embryonic SC differentiation by controlling the dosage of β -catenin signaling. Moreover, LGR5 that is overexpressed in human CRCs (66) is a Tcf4 target gene (70) and LGR5 was then used to identify crypt SCs as the cells-of-origin of intestinal cancer (71). Taken together, these findings show that *APC* mutations and activation of WNT signaling pathways are crucial to the development of CRC.

APC AND WNT GRADIENTS

How crypt SC overpopulation is caused by *APC* mutations remains unclear. An explanation we are presenting here is that, in

the normal crypt, APC-induced down-regulation of WNT signaling establishes an APC:WNT gradient and dysregulation of this gradient in tissues containing *APC* mutations is key. In normal crypt renewal, daughter cells produced by SCs at the bottom of the crypt proliferate while they migrate upwards. Because APC protein produced in crypt cells increases as the cells migrate upwards (58, 72–80), APC concentrations are low at the bottom of the crypt (where SCs reside) and high at the top of the crypt (where differentiated cells are). In contrast, WNT signaling is greater at the bottom of the crypt, occurring through a complex network consisting of different WNT ligand and receptor signaling components (81, 82). Activation of WNT signaling in the crypt bottom was shown by studies demonstrating accumulation of nuclear TCF4 in the crypt proliferative compartment (70, 83–85). There are several lines of evidence demonstrating that continual stimulation of the WNT pathway in the crypt bottom is essential for maintenance of intestinal SCs, normal proliferation of transit-amplifying cells, enterocyte maturation, and crypt homeostasis (86, 87). WNT gradients are high at the crypt bottom and low at the crypt top (inverse to the APC gradient). Given the existence of these inverse gradients and the dynamics of their interactions, we construed that there is a counter-current-like mechanism in the normal crypt and that this mechanism likely regulates changes in cellular phenotype associated with colonocyte maturation along the crypt axis (Figure 1).

COUNTER-CURRENT-LIKE MECHANISMS

Counter-current mechanisms are found extensively in nature. Typically, the incoming and outgoing components flow in opposite directions to each other and interact to retain a high concentration of a substance at one point in the system. In the colonic crypt, one component is the Wnt gradient and the other is the APC gradient. These gradients are not only “counter” to each other, but also APC and Wnt are both necessary for proliferation (but neither is sufficient). As discussed below, APC and Wnt components are known to interact during mitosis and, in our model, this interaction maintains cell division at a high level in a specific area in the lower region of the normal crypt (peak at approximately crypt level 15).

This raises the question as to what maintains the APC and Wnt gradients in the normal colonic crypt. The inverse pattern is consistent with their being feedback and/or feed-forward regulation, which is a key to many counter-current mechanisms. For example, Wnt signaling is activated at the crypt bottom by a complex network of various Wnt ligands and receptors. The Wnt and APC gradients can even affect each other. Indeed, it is well-known that APC down-regulates Wnt signaling.

REGULATION OF APC EXPRESSION

Of course the question arises: What factors regulate *APC* gene expression and is WNT signaling one of those factors? One factor appears to be cell proliferation. For example, Umar et al. (75) found that epithelial proliferation induces APC expression and full-length APC protein increases during rodent intestinal epithelial hyper-proliferation. Fagman et al. (88) showed a similar effect. Their study showed that nuclear accumulation of full-length and truncated APC protein in colon carcinoma cell lines depends on

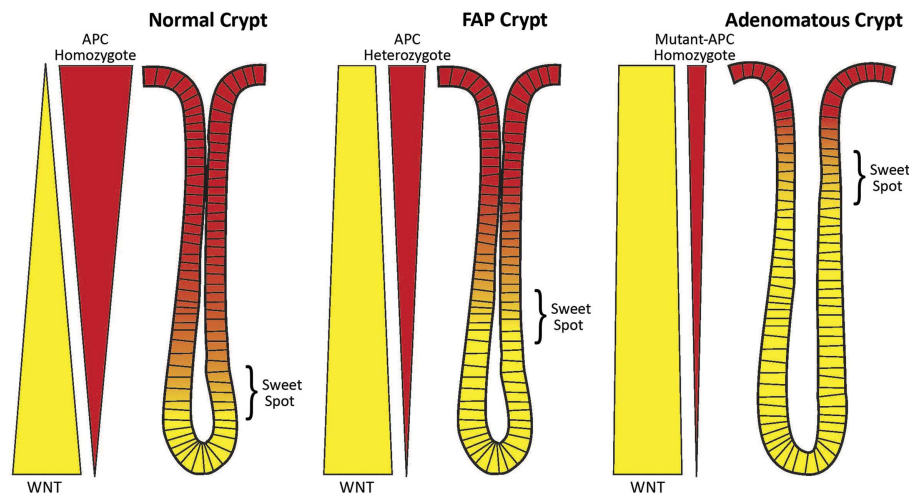


FIGURE 1 | Schematic of the counter-current-like mechanism. The “sweet spot” marks the crypt region where levels of APC and WNT signaling are balanced and optimal conditions exist for mitosis and maximal cell proliferation. Both APC and WNT signaling components (e.g., survivin) are essential for mitosis. Left Panel: In normal crypts (wt-APC homozygote), the gradients of WNT signaling (yellow wedge) and APC signaling (red wedge) are balanced and the “sweet spot” is in the lower crypt. Middle Panel: In FAP crypts (wt-APC heterozygote) the situation has changed. Patients with FAP have a germline heterozygous APC mutation and thus a 50% reduction in APC gene dosage. Therefore, there is 50% less APC protein expressed (as indicated by the narrower red wedge), and less

suppression of WNT expression and WNT signaling (as indicated by the wider yellow wedge). The balance point, that is the “sweet spot,” has been shifted to a higher crypt level. Right Panel: In adenomatous crypts (mutant-APC homozygote), the changes in WNT expression (still wider yellow wedge) and APC expression (still narrower red wedge) are exacerbated due to a sporadic APC mutation in the second APC allele (the second hit). In mutant-APC homozygote cases truncated APC protein can retain some residual function. Here the “sweet spot” is shifted even further up the crypt. A consequence of these changes is an increase in the number of immature cells (including SCs) in the crypt. The SC overpopulation is thought to drive colon tumorigenesis.

proliferation. Another factor is that regulation of APC expression depends upon promoter methylation. Deng et al. (89) showed that methylation of CpG sites around a CCAAT box in APC's promoter region inhibits APC's gene expression by changing chromatin conformation and interfering with the binding of transcription factor CBF (CCAAT binding factor) to the CCAAT box. Studies on various cancers provide further support for the idea that expression of APC is affected by promoter methylation. Indeed, APC promoter hyper-methylation has been found to occur in a variety of human cancers including breast (44%) and lung (53%) and other cancers (90). This promoter hyper-methylation leads to epigenetic inactivation of the APC gene. Some transcription factors such as p53, USF1, USF2, and GC-box binding protein Sp3 have also been shown to regulate APC gene expression (91–95). While it has not been shown (to our knowledge) that Wnt signaling directly regulates APC expression (APC is not a known TCF4 target gene), WNT signaling does have a role in chromatin remodeling via beta-catenin's interaction with chromatin remodeling complexes that can affect gene transcription (96). This role fits with the observation that chromatin is more condensed in cells at the crypt base, the region where WNT has highest activity (97). Thus, it is changes in chromatin conformation (open or closed) along the crypt axis that may modulate APC expression by affecting the ability of specific transcription factors to bind to APC's promoter region.

APC AND WNT SIGNALING ARE BOTH REQUIRED FOR MITOSIS

In particular, the existence of inverse APC and WNT gradients and their interactions begins to explain how crypt cell dynamics such

as cell division are regulated. For instance, APC is a protein located at specific sites within mitotic cells and is essential for cell division [reviewed in (98, 99)]. Cell division also requires WNT signaling. In the WNT signaling pathway key down-stream signaling components include survivin, aurora B kinase (ABK), INCENP, and phospho-histone H3 (58, 100). ABK becomes activated when it forms a complex with the other three proteins. Like APC, ABK, Survivin, INCENP, and Borealin are also located at specific sites within mitotic cells and these proteins are necessary for cell division (101–103). We and others found that survivin is a TCF4 target gene (58, 63, 100, 104–106). Thus, APC itself down-regulates, via beta-catenin/TCF4, expression of survivin. This, in turn, modulates ABK activity, which contributes to the regulation of mitosis.

As both APC and WNT signaling are essential for mitosis, and as their gradients are inverse to one another, there has to be a zone along the crypt axis where the concentrations of APC and WNT pathway components are together optimal for maximal cell division – termed here a “sweet spot.” To better understand the underlying mechanisms of this phenomenon, we created a new model. It is a counter-current-like model that considers gradients of factors (APC; WNT) that regulate colonocyte proliferation along the crypt axis (Figure 1). The scientific basis of this model derives from the fact that both APC and WNT signaling are both required for mitosis, which we will now discuss.

ROLE OF APC IN SPINDLE ORIENTATION AND MICROTUBULE FUNCTION DURING MITOSIS

During mitosis, APC becomes localized to and acts at four sub-cellular sites: midbody, centrosomes, cortex, and kinetochores

[reviewed in (98, 99)]. APC proteins have several mitosis-related functions. (a) APC acts at the plus ends of microtubules, which, in turn, interact with kinetochores. This increases kinetochore ability to attach to microtubules. APC also helps maintain mitotic fidelity – APC is needed for the normal function of the mitotic spindle checkpoint, which includes detection of transiently misaligned chromosomes. (b) At the cortex, APC regulates the stability of astral microtubules, and provides a cortical location for attachment of those microtubules. APC and other interacting cortical factors (which includes dynein) rotate the mitotic spindle into a defined orientation, and help orient the mitotic spindle. (c) In vertebrates, APC and EB1 (its binding partner) co-localize to the mother centriole, and this anchors a subset of microtubules. At this site, it is likely that APC has a role in centrosome re-orientation and directed migration. (d) APC also co-stains with tubulin at the midbody (which separates daughter cells) suggesting a role for APC in cytokinesis (107).

APC is associated, during mitosis, with kinetochores in metaphase, with the midbody during telophase and with polar microtubules in anaphase. During metaphase, APC interacts with EB1 in the alignment of chromosomes through plus-end capture and through attachment of microtubules to kinetochores (108, 109). This linkage seems to require EB1 and, notably, APC and EB1 localize specifically to the mother centriole. APC is also localized to the centrosome, and helps nucleate and anchor microtubules, which is required for establishment of the bipolar spindle.

ROLE OF WNT SIGNALING IN MITOSIS AND MITOTIC SPINDLE ORIENTATION

Survivin, ABK, INCENP, and Borealin are also involved in mitosis and function as a complex of chromosomal passenger proteins [chromosomal passenger complexes (CPCs)] localized to chromosomes in prophase (101–103). In metaphase, survivin targets this CPC to centromeres. The complex is stabilized by INCENP. Survivin binding to INCENP is promoted by Borealin. During metaphase, ABK is localized to the inner centromere. ABK regulates kinetochore-microtubule interactions and promotes proper chromosome bi-orientation by regulating and correcting kinetochore-microtubule attachments. In particular, ABK at the inner centromere inhibits formation of syntelic microtubule attachments, thus promoting monotelic attachments and appropriate bi-orientation on the mitotic spindle.

APC, EB1, SURVIVIN, AND ABK LOCALIZE TO SIMILAR SITES DURING THE DIFFERENT PHASES OF MITOSIS

The cellular location of APC, EB1, survivin, and ABK during the different phases of mitosis has been described in several reviews (98, 99, 102).

In anaphase APC is located at the cortex and helps position the spindle. This depends on guiding cortical microtubules to specific cortical sites. This seems to require a microtubule plus-end protein complex, which includes APC, beta-catenin, EB1, and other proteins. APC is linked to microtubules by binding to EB1, and dynactin/dynein complexes are tethered to APC-associated EB1 at cortical attachment sites. Indirect links between APC and actin filaments seem to be mediated by beta-catenin and alpha-catenin, which provides, during mitosis, a functional link between

microtubule and actin cytoskeletons. Survivin and ABK concentrate, during anaphase, at the spindle midzone and equatorial cortex in preparation for their roles in late mitotic events.

In telophase, APC co-stains with tubulin at the midbody (107). It is unknown how APC contributes to furrow induction during cytokinesis. But it may help guide cortical microtubules to the cortex, or control actin dynamics at the cortex. ABK and survivin also play critical roles in cytokinesis. During telophase, ABK and survivin are localized to the midbody. ABK seems to mediate cytokinesis by phosphorylating several proteins that localize to the cleavage furrow, which destabilizes intermediate filaments prior to cytokinesis (110, 111).

DO APC, SURVIVIN, AND ABK INTERACT DIRECTLY?

The fact that APC, EB1, survivin, and ABK localize to similar sites during the different phases of mitosis suggests that interactions occur between these proteins. Indeed, the APC binding protein EB1 provides a link between APC and ABK because EB1 and ABK co-localize to the central spindle in anaphase and to the midbody during cytokinesis. For instance, it was found that EB1 promotes ABK activity by blocking its inactivation by protein phosphatase 2a (112). Therefore, EB1 mediates microtubule dynamics in association with APC and also positively regulates ABK activity. In addition, formin mDia3, another APC binding protein, helps stabilize microtubule-kinetochore attachments and chromosome alignment in metaphase (113). This ability has been attributed to the binding of mDia3 to EB1, the other protein that interacts with APC. This provides another link to ABK because microtubule binding to kinetochores via mDia3 is regulated by ABK phosphorylation of mDia3 (114). Thus, during mitosis, there are both anatomical and functional links between APC, EB1, and ABK.

The evidence thus indicates that these key mitotic components need complex regulation during the different phases of mitosis and at different locations along the crypt axis. One aspect of this regulation, we and others found, involves APC itself. APC down-regulates expression of survivin via beta-catenin/TCF4, which, in turn, modulates ABK activity (58, 63, 100, 104–106). APC, which is a tumor suppressor gene, not only helps in mitosis but also promotes both differentiation and apoptosis in the colonic crypt. As noted above, APC in the crypt is distributed along a gradient, from essentially negligible at the crypt bottom to a maximum at the crypt top. The WNT gradient is inverse to the APC gradient. Since survivin is a down-stream component of WNT signaling, and survivin activates ABK, it is not surprising that survivin and ABK gradients are, like the WNT gradient, highest at the crypt bottom and lowest at the crypt top [(58, 100)].

That these gradients are inverse to one another might be seen as contradicting the fact that APC, ABK, and survivin are essential for appropriate progression of cells through mitosis. However, what this evidence really provides is insight into how cells undergo phenotypic transitions as they migrate and undergo maturation upwards along the crypt axis. In that view, there is, along the normal crypt axis, a region where APC, EB1, survivin, and ABK levels together generate the highest rates of proliferation and differentiation. In this “sweet spot,” transit-amplifying cells proliferate and differentiate – a concept that is supported by our mathematical modeling and biologic data (23, 58, 62, 63).

OUR COUNTER-CURRENT-LIKE MECHANISM IS A MODEL THAT ALSO EXPLAINS THE PROLIFERATIVE CHANGES DURING COLON TUMORIGENESIS

In heterozygous *APC* mutant crypts, such as normal-appearing colonic epithelium in FAP patients, there is half the wild-type *APC* gene dosage. But both *APC* alleles will still be regulated as in the normal crypt. However, one of the transcripts will be a mutant transcript and be translated into a mutant protein (50%) while the other will encode a wild-type protein (50%). Thus, the encoded wild-type protein levels should be reduced by about half at all points along the haplo-insufficient crypt axis and the gradient becomes diminished but retains its distribution. Because concentrations of wt *APC* are decreased, the optimal concentrations that generate the sweet spot occur further up the crypt (**Figure 1**). This upward shift in the position of the sweet spot parallels the shift in the labeling index in FAP crypts from the lower region to the middle crypt (see Proliferative Changes in *APC* Mutant Crypts During Human Colon Tumorigenesis, above).

In homozygous mutant *APC* crypts, such as are found in adenomas, these changes become exacerbated further. In this case, the *APC* gradient is not totally lost because the truncated *APC* protein usually contains some residual function (48, 49). Thus, optimal conditions corresponding to the sweet spot are found even further up the crypt. Indeed, the labeling index in adenomatous crypts shifts to the top of the crypt.

RESIDUAL APC FUNCTION IS RETAINED IN CELLS WITH HOMOZYGOUS MUTANT APC

It is fascinating how *APC* mutation leads to retention of residual activity in the encoded mutant protein. In the situation where tumors are homozygous mutant for *APC*, the site of the “first hit” in the *APC* gene determines the type of the “second hit,” both in hereditary (FAP) and sporadic colorectal tumors (48, 115–118). This results in expression of truncated *APC* protein in most tissues with homozygous mutant *APC*. But, the truncated *APC* protein actually retains a microtubule binding domain (Armadillo repeats) and one to three intact β -catenin-binding amino acid repeats. This indicates that second hits at the *APC* locus occur that generate a “just-right” level of WNT/ β -catenin signaling that is optimal for tumorigenesis, with the combined hits (or “just-right” genotypes) resulting in only partial loss of *APC* functioning (49, 119, 120).

Therefore, in neoplastic crypts one has to look higher up the crypt to find the optimal *APC* concentration for cell division. This is also true for WNT because diminished *APC* leads to diminished down-regulation of WNT. Thus, it is only higher up the crypt where the *APC* concentration is high enough to diminish WNT levels to what normally was the optimal WNT concentration for promoting cell division. Shifting the proliferative zone upwards in neoplasia will theoretically increase symmetric SC division below the sweet spot, which will cause CSC overpopulation and promote colon tumorigenesis. Mutations at the second *APC* allele would exacerbate these changes. Thus, *APC*-mutation-induced changes in a counter-current-like mechanism will increase the number of proliferative cells (SCs, rapidly proliferating cells), contributing to colon cancer initiation and adenoma development.

Indeed, FAP crypts have increased mitoses (61), and a cardinal pathologic feature of colonic adenomas and carcinomas is increased numbers of mitotic figures and aneuploidy. One explanation for this increase is that the residual *APC* activity in combination with increased WNT signaling in neoplastic tissues increases the frequency of mitosis. If *APC* is necessary (but not sufficient) for mitosis, and if there is enough residual *APC* function in a tumor, one would expect more frequent mitoses (but not a greater rate of mitosis) when WNT signaling is also upregulated. In the setting where you have increased frequency of mitosis, you could also have changes in the fidelity of chromosome segregation. It is probably the dynamic interplay between *APC* and CPC proteins during mitosis that affects the fidelity of mitosis (e.g., the accurate segregation of chromosomes). Since the truncated *APC* protein is not fully functional, and leads to, via WNT, aberrant down-regulation of CPC protein expression, it is not surprising that *APC* mutations lead, in colonic tumors, to aberrantly oriented mitotic spindles and aneuploidy.

EFFECTS OF CHANGES IN COLONIC APC AND WNT GRADIENTS IN NEOPLASTIC CRYPTS ON SYMMETRIC AND ASYMMETRIC CELL DIVISION IN THE SC NICHE

The counter-current-like mechanism may also play a role in regulating the symmetry of crypt cell divisions. The orientation of the mitotic spindle axis of colon cells appears to change upward along the normal crypt axis. In the crypt bottom, the mitotic spindle orientation lies perpendicular to the apical surface – an orientation that selectively occurs in the SC niche of human and rodent small and large intestine (121). This perpendicular alignment of mitotic spindles correlated with the pattern of retention of label-retaining DNA in the crypt base, a pattern that is consistent with asymmetric division of SCs. In the normal crypt middle (i.e., in the crypt column), where SCs are rarely found, Quyn et al. (121) showed that mitotic spindle alignments are mostly parallel to the apical surface, which is a pattern that is consistent with symmetric differentiated cell division. This mechanism involving change in the mitotic spindle orientation along the normal crypt axis could contribute to the maintenance of a constant number of SCs in the SC niche (16, 64).

In premalignant heterozygous mutant *Apc* tissue in rodents (*Apc*^{Min/+} intestine), both perpendicular spindle orientations and asymmetric DNA label retention were lost in the SC niche (but not in the crypt middle) (121). This pattern is consistent with a decrease in asymmetric and an increase in symmetric cell division in the SC niche in crypts with an *APC* mutation. Moreover, their murine small intestine and human colon data demonstrate that crypts with *APC* mutations show increased asymmetric cell division in the crypt middle (column). These data thus indicate that a shift in asymmetric cell division from the crypt bottom to the crypt column happens in parallel to the shift in the labeling index that was reported by others (20–22, 50, 56, 57). We also reported (58) that the subpopulation of cells expressing the mitotic proteins ABK and survivin shifts upward in FAP and *Apc*^{Min/+} crypts. The global effect on the crypt of an *APC* mutation thus is a delay in phenotypic transitions along the crypt axis, an increase in the number of SCs that divide symmetrically, and expansion of the SC population at the crypt bottom, which drives colon tumorigenesis.

STUDIES ON RODENT INTESTINE THAT MIGHT PROVIDE INSIGHT INTO MECHANISMS OF SC DIVISION IN HUMAN COLON

Theoretically, in normal tissue renewal, asymmetric cell division maintains the number of SCs constant (37, 122). An alternative concept is that SCs must, on average, have asymmetric divisions, even if each particular SC division is not always asymmetric (123). In neoplastic tissue, in contrast, it can be deduced that, in the development of SC overpopulation during human colon tumorigenesis, the rate of symmetric SC division must be increased (37). However, it is controversial whether, in the process of crypt renewal, SCs normally divide symmetrically or asymmetrically. Yatabe et al. (124) used methylation patterns to investigate this question for human colonic SCs. The patterns better supported a model of the crypt SC niche in which SCs were periodically replaced via symmetric SC division. And, using this methylation pattern diversity analysis, Kim et al. (65) found enhanced SC survival in FAP, which is consistent with SC overpopulation in CRC development. In contrast, a recent paper by Bu et al. (34) reports data showing that human colon cancer SCs can divide by either symmetric or asymmetric division.

There is now a vast literature emerging on mechanisms of SC division in rodent intestine that provide insight into mechanisms of SC division in human colonic crypts. As noted above, *Lgr5* was identified to be a SC marker in mouse intestine (67). Using lineage-tracing models that were based on fate mapping, Snipert et al. (125) reported that rapidly cycling small bowel SCs in rodents (i.e., *Lgr5*+ cells), undergo symmetric SC division that follows a pattern of “neutral drift dynamics.” These findings support a stochastic mechanism in which symmetric SC division occurs in response to loss of a neighboring SC and, as expansion of the surviving clone continues, the crypt SC niche becomes increasingly monoclonal (126). Schepers et al. (127) studied SCs (*Lgr5*+ cells) in the base of the crypt looking for asymmetric segregation of chromosomes and asymmetric segregation of chromosomes was not observed at the crypt base since *Lgr5*+ intestinal SCs randomly segregated newly synthesized DNA strands.

Other studies, however, indicate that SCs divide asymmetrically. Asymmetric division of SCs in the SC niche supports a deterministic mechanism whereby a small number of SCs each generate a SC and a non-SC (a transit-amplifying cell). The non-SC daughter leaves the SC niche, and proliferates (promoting tissue renewal) whereas the SC daughter stays in the SC niche (35, 128). Asymmetric SC division is consistent with the “immortal strand hypothesis,” that is, the idea that during SC division, newly synthesized DNA strands segregate with the non-SC daughter to avoid mutations that are caused during DNA replication (129).

The “immortal strand” hypothesis has been tested using DNA-labeling methods to identify label-retaining cells (LRCs). Cells that retain DNA labels like BrdU or [³H]thymidine are thought to be SCs. Potten et al. (130) used double-labeling of cells in rodent small bowel using BrdU and [³H]thymidine (³HTdR). Template DNA strands in SCs were labeled with ³HTdR during tissue regeneration or development. Newly synthesized strands were labeled with BrdU, which established a way to follow how the two markers segregated after cell division. The authors found that the template strands (which were ³HTdR-labeled) were retained,

but newly synthesized strands (which were BrdU-labeled) were lost following the second SC division. Studying cultured cells that cycle with asymmetric cell kinetics, Merok et al. (131) reported co-segregation of chromosomes containing immortal DNA strands, and that is also consistent with the immortal strand hypothesis. Moreover, as discussed above, Quyn et al. (121) observed that labeled DNA was asymmetrically retained in the SC niche of rodent intestinal crypts. The pattern of retention of label-retaining DNA correlated with the perpendicular alignment (alignment relative to the apical surface) of mitotic spindles in the crypt base.

In a study on rodent colon, Kim et al. (65) showed that a double-labeling method (BrdU and ²H₂O) could be employed to identify and isolate nuclei from colonic epithelial LRCs. This let them measure proliferation rates of LRCs *in vivo* ($t_{1/2} \approx 140$ days). Falconer et al. (132) used fluorescence *in situ* hybridization and unidirectional probes specific for centromeric and telomeric repeats. They found that one can identify parental DNA template strands in sister chromatids of rodent metaphase chromosomes. These findings showed that orientation of chromosomes is uniform; the 5' end of the short arm is on the same strand as the “T-rich” major satellite repeats. This repetitive DNA orientation allows both analysis of mitotic segregation patterns and differential labeling of sister chromatids. The authors uncovered substantial non-random segregation of sister chromatids in a subpopulation of colonic crypt epithelial cells, which included cells outside the SC niche. This finding suggested that there exists in colonocytes a mechanism that controls how sister chromatids are allocated as intestinal SCs divide. Interestingly, DNA methylation is emerging as a mechanism that might regulate non-random template strand segregation suggesting that this aspect of SC division may be under dynamic control.

How can one reconcile these different findings regarding symmetric SC division and asymmetric SC division in the population of intestinal SCs during crypt renewal? One possible way is based on recent findings that, in the SC niche, different subpopulations of SCs exist. Using a variety of markers (*Bmi1*, *Lgr5*, and *mTert*), distinct intestinal SC populations were identified (133). *Lgr5* labels SCs that rapidly cycle and are located in the crypt base, modulated by Wnt signaling, and sensitive to irradiation (71). Subsequently, using lineage-tracing experiments in adult rodents, Barker et al. (134) showed that cycling *Lgr5*+ cells are very long-term self-renewing cells in the intestine. Another study identified *Lrig1* protein, the pan-ErbB negative regulator, as a specific intestinal SC marker which also functions as a tumor suppressor (135). In a different study, Montgomery et al. (136) identified a subpopulation of intestinal SCs that express telomerase reverse transcriptase (*mTert*), that cycle slowly, and that give rise to *Lgr5*+ cells. This study showed that although *Lgr5*+ intestinal SCs represent a different subpopulation, they can also have high telomerase activity. *Bmi1* labels a different subpopulation of intestinal SCs. These SCs are quiescent, insensitive to Wnt signaling, resistant to high-dose radiation, and generate all the differentiated lineages in the crypt. Since *Bmi1* and *Lgr5* label two different populations of SCs and since *Bmi1*+ cells can generate *Lgr5*+ cells, *Bmi1*+ cells represent a reserve SC population that cause *Lgr5*+ cells to be dispensable (137). Thus, based on the above findings, the idea that SCs are quiescent has been challenged. Some studies suggest that there

are slowly cycling SCs; other studies suggest that there are actively cycling SCs; yet other studies suggest that there are only relatively quickly cycling SCs (138–140). This evidence demonstrates that there may be different SC subpopulations in the intestinal SC niche and suggests that these different populations may even have different modes of division (asymmetric SC division vs. symmetric SC division).

Another possible explanation is provided in a recent editorial by Winton (141) in which he states “Interpretation of the growth dynamics of stem-cell-derived clones has previously demonstrated that symmetric fate choice is a common feature of intestinal SC self-renewal (126). However, asymmetric fate choices could be interspersed with symmetric ones and still be compatible with these models [e.g., if small numbers of SC per crypt are assumed; (125, 142)].” Another possible explanation is given by a recent study by Bellis et al. (143). These investigators found that *Apc* controls planar cell polarities that are central to gut homeostasis. By studying the SCs at the bottom of intestinal crypts, they discovered that spindle alignment and planar cell polarities form a functional unit that can generate daughter cell anisotropic movement away from niche-supporting cells. By proposing a mechanism involving anisotropic daughter cell movement rather than spindle re-orientation in SCs [per (121)], the Bellis model provides an alternate mechanism for the idea of neutral competition of SCs for niche-supporting cells, that is central to the concept of stochastic population asymmetry (128).

Alternatively, there are other studies that may reconcile the ambiguity as to whether SCs are rapidly cycling or quiescent. Buczacki et al. (24) found that quiescent intestinal cells in rodents are precursors that are committed to maturing into differentiated secretory cells of the Paneth and endocrine lineages. However, upon intestinal injury, they become capable of extensive proliferation and give rise to the other intestinal cell lineages. Thus, quiescent intestinal crypt cells represent a reserve population that can be recruited to a SC state. A study by Takeda et al. (43) also showed that there is inter-conversion between intestinal SC populations (Hopx+ slow cycling SCs and Lgr+ proliferating SCs) in distinct niches. Kobayashi et al. (144) reported that Lgr5+ colon cancer SCs interconvert with drug-resistant Lgr5– cells, which are capable of tumor initiation. Glauche et al. (145) reported that SC proliferation and quiescence were two sides of the same coin. They concluded that “hematopoietic SC organization was an adaptive, regulated process where the slow activation of quiescent cells and their possible return into quiescence after division are sufficient to explain the simultaneity of occurrence of self-renewal and differentiation.”

Clearly, a great deal of research has been done to study the various intestinal SC populations (e.g., Lgr5, Hopx, mTert, Bmi1, Lrig1, etc). However, the interacting dynamics and modes of division of these different intestinal SC types do not appear to be fully resolved. Taken together, we believe that these studies suggest that the intestinal SC is a cell that is in one of several phenotypic states that immature enterocytes can assume based on the dynamics within the colonic crypt. These dynamics may be relevant to those proposed in our counter-current-like mechanism proposed for human colonic crypts.

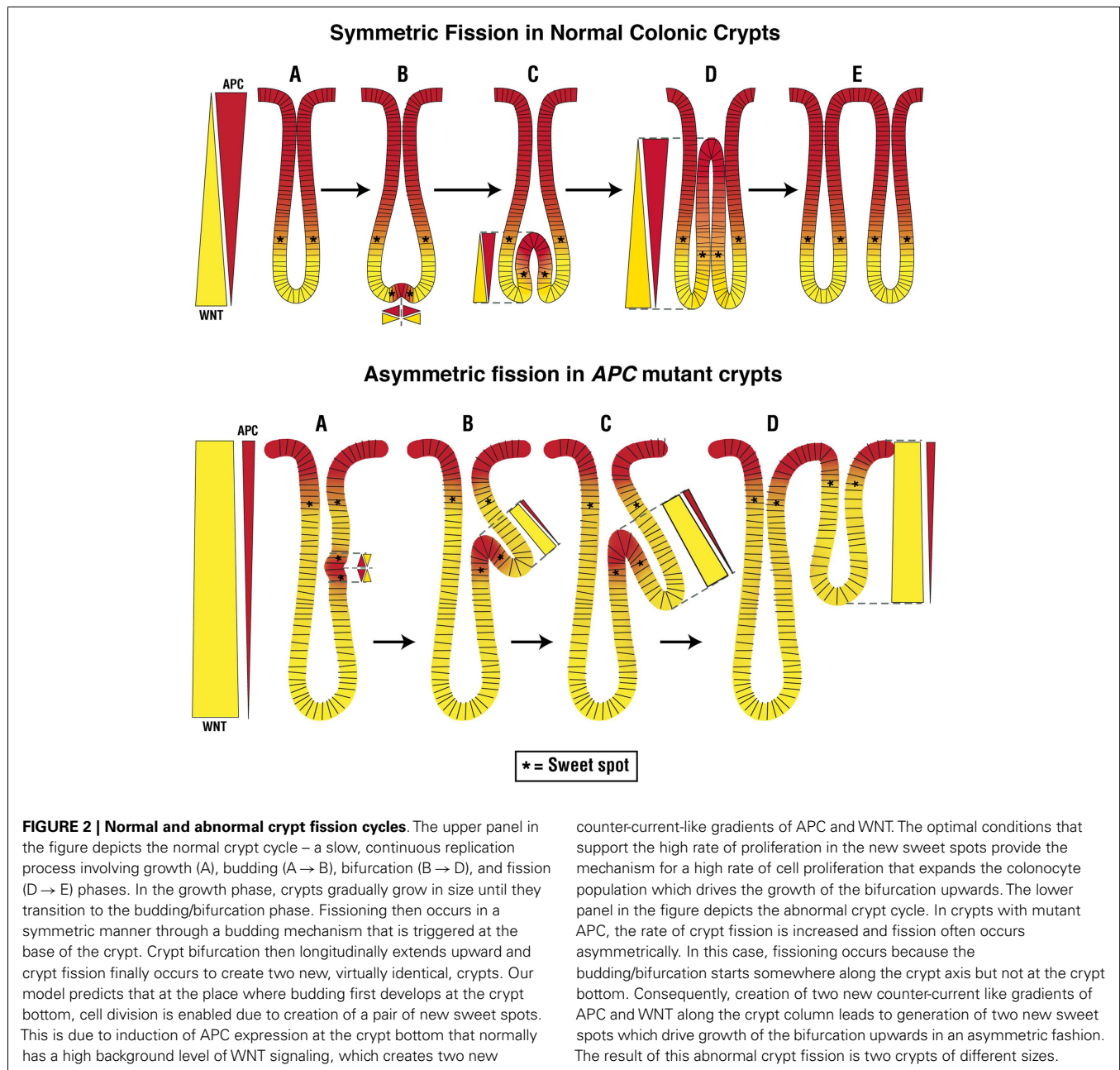
COUNTER-CURRENT-LIKE MECHANISMS AND ADENOMA DEVELOPMENT

Adenoma morphogenesis is due, in large part, to abnormal crypt fission. In the normal adult rodent intestine and human colonic crypt, fission is responsible for regular replacement of crypts through the “crypt cycle” (146–150). The crypt cycle is a slow, continuous replication process involving three phases (growth, budding/bifurcation, and fission) (**Figure 2**, upper panel). In the growth phase, crypts gradually grow in size until the transition to the budding/bifurcation phase. The fissioning process then occurs in a symmetric manner through a budding mechanism that is triggered by a development of a bud (appearing as an indentation) at the base of the crypt. Crypt bifurcation then longitudinally grows and extends upward and crypt fission finally occurs to create two new virtually identical crypts (60, 151, 152). Two factors have been proposed to govern crypt fission: the size of the crypt and the size of the crypt SC population (147, 153). Since the crypt cycle produces expansion in the crypt population size, it is critical for epithelial homeostasis, as well as for repair of mucosal injury.

In adenomas that develop due to *APC* mutations, tissue disorganization is manifest as dysplasia, and premalignant tumor growth results from an increased rate of intestinal crypt fission (151, 154–156). For example, Wasan et al. (60) showed that both FAP patients and *Apc*^{Min/+} mice have increased rates of intestinal crypt fission in *APC* haplo-insufficient intestinal epithelia. In homozygote *APC* mutant epithelium, the rate of crypt fission is even greater (153). This identifies *APC* as one of the key factors in the regulation of crypt fission. Moreover, an increase in the crypt fission rate appears to account for the clonal and exponential expansion of mutant cell populations that drive tumor growth (157, 158). In normal-appearing and adenomatous intestinal tissues from FAP patients and *Apc*^{Min/+} mice, histologically aberrant crypt fissioning occurs. In this process, the budding/bifurcation process is asymmetrical, giving rise to crypt branching and non-identical crypts (60, 153).

So how might these changes in crypt fission relate to our counter-current-like model? In normal crypts, budding initiates the fissioning process at the bottom of the crypt. In our model (**Figure 2**), fissioning normally begins below the sweet spot at a point in the SC niche where WNT signaling is highest and *APC* is lowest. In this scenario, the *APC*/WNT gradients restrict the region where fissioning can be initiated to the crypt bottom, such that fissioning will proceed symmetrically.

In heterozygote and homozygote *APC* mutant crypts, asymmetric crypt fissioning appears to occur because fissioning starts anywhere along the crypt axis, not just at the crypt bottom (60, 159). In our model (**Figure 2**), this happens in a crypt that has an upward shift in the sweet spot and an expansion of the region below the sweet spot where low *APC* and high WNT levels occur. In this view, a change in *APC* and WNT gradients expands the region where fissioning can be initiated so it can occur anywhere along the crypt column including toward the crypt top such that fissioning occurs asymmetrically. But, in *APC* mutant crypts, not only is the symmetry of fissioning abnormal, but the rate of fissioning is increased. Based on our counter-current-like mechanism, changes in *APC* and WNT gradients due to *APC* mutation lead to



an increased WNT gradient (while the APC gradient diminishes). In this view, increased WNT signaling not only increases the rate of crypt fissioning but also causes asymmetric fissioning, which underlies adenoma growth. This view is consonant with studies implicating Wnt/ β -catenin signaling in crypt fission because WNT is essential for intestinal SC division (152, 160).

To further understand how our counter-current-like mechanism might relate to aberrant crypt fissioning (asymmetric fissioning and increased rate of fissioning) that drives adenoma development, it is useful to draw parallels between the crypt cycle and the hair follicle cycle. In the hair follicle cycle, hair grows cyclically through three phases: *anagen* is the growth phase; *catagen* the involuting or regressing phase; and *telogen*, the resting or quiescent

phase (161). As noted above, there are also three phases in the crypt cycle: the crypt *growth* phase, the crypt *budding/bifurcation* phase, and the crypt *fission* phase. In the hair follicle cycle, WNT signaling maintains the anagen growth phase (162). That WNT signaling and the rate of crypt fission are both increased in APC mutant crypts suggests that WNT signaling also has a role in the growth phase of the crypt cycle.

Since, based on our model, “optimal” APC levels in mutant crypts occur higher up the crypt and fissioning occurs asymmetrically at points higher along the crypt axis, APC may also have a role in the budding/bifurcation phase of the crypt cycle. Indeed, it has been proposed that it is APC that normally controls the symmetry of crypt fissioning (60). Our model predicts that at the place

where budding first develops at the crypt bottom, cell division is enabled due to localized induction of APC expression that establishes two new gradients which creates a pair of new sweet spots (Figure 2). The enabling of rapid cell division at these new sweet spots creates a motor mechanism that drives growth of the bifurcation upwards toward the crypt top. This proposed mechanism is consonant with biological data showing that increased cell division selectively occurs on both sides of the extending bifurcation in fissioning crypts (163). Moreover, cells staining positively for the Wnt target gene *Lgr5* are located at the bottom-most region of the two newly emerging crypts (164). These *Lgr5*⁺ cells in bifid crypts appear to be located below both sides of the extending bifurcation. Thus, our model predicts that optimal APC and WNT signaling are crucial to regulating the rate and symmetry of crypt fissioning during the crypt cycle. Therefore, based on our mechanism, changes in APC and WNT that are due to APC mutation alter regulation of the crypt cycle, cause abnormal crypt fission, and drive adenoma development.

CLINICAL SIGNIFICANCE

Based on our proposed counter-current-like mechanism, it may be possible to develop novel approaches that normalize the APC and

WNT gradients, shift the proliferative zone downwards, and thwart the progression of premalignant changes in the APC mutant colonic crypt. One can consider targeting APC, but it is unlikely that gene therapy will be efficient enough to transfect wt-APC genes into mutant SCs. The alternative is to diminish the WNT gradient. In principle, this can be done by targeting TCF4 or TCF4 target genes such as survivin. Indeed, several agents that inhibit TCF4 activity are already in development (165, 166).

CONCLUSION

Our consideration of how APC mutations affect the spatial and temporal organization of the colonic crypt led us to propose an APC:WNT counter-current-like mechanism that regulates cell division along the crypt axis. It is a mechanism that explains how APC mutations induce proliferative abnormalities that drive colon cancer development. This mechanism also suggests how chemoprevention for this malignancy might be achieved.

ACKNOWLEDGMENTS

Generous support for this work was provided by the Helen F. Graham Cancer Center, the Bioscience Center for Advanced Technology (CAT) and the Cancer B Ware Organization.

REFERENCES

- Kennell J, Cadigan KM. APC and beta-catenin degradation. In: Nathke IS, McCartney BM, editors. *APC Proteins*. New York, NY: Springer Science+Business Media, LLC (2009). p. 1–12.
- Polakis P. Wnt signaling and cancer. *Genes Dev* (2000) 14:1837–51.
- Clevers H. Stem cells: a unifying theory for the crypt. *Nature* (2013) 495:53–4. doi:10.1038/nature11958
- Schuijers J, Clevers H. Adult mammalian stem cells. The role of WNT, *Lgr5* and R-spondins. *EMBO J* (2012) 31:2685–96. doi:10.1038/emboj.2012.149
- Scoville DH, Sato T, He XC, Li L. Current view: intestinal stem cells and signaling. *Gastroenterology* (2008) 134:849–64. doi:10.1053/j.gastro.2008.01.079
- Potten CS, Kellett M, Roberts SA, Rew DA, Wilson GD. Measurement of in vivo proliferation in human colorectal mucosa using bromodeoxyuridine. *Gut* (1992) 33:71–8. doi:10.1136/gut.33.1.71
- Arai T, Kino I. Morphometrical and cell kinetic studies with normal human colorectal mucosa: comparison between the proximal and distal large intestine. *Acta Pathol Jpn* (1989) 39:725–30. doi:10.1111/j.1440-1827.1989.tb02421.x
- Cheng H, Bjerknes M, Amar J. Methods for the determination of epithelial cell kinetic parameters of human colonic epithelium isolated from surgical and biopsy specimens. *Gastroenterology* (1984) 86:78–85.
- Cheng H, Leblond CP. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. I. Columnar cell. *Am J Anat* (1974) 141:461–79. doi:10.1002/aja.1001410404
- Cheng H. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine II: mucous cells. *Am J Anat* (1974) 141:481–502. doi:10.1002/aja.1001410404
- Shamsuddin AM, Phelps PC, Trump B. Human large intestinal epithelium: light microscopy, histochemistry and ultrastructure. *Hum Pathol* (1982) 13:790–803. doi:10.1016/S0046-8177(82)80075-0
- Lorenzson V, Trier JS. The fine structure of human rectal mucosa. *Gastroenterology* (1968) 55:88–101.
- Winton DJ, Ponder BA. Stem-cell organization in mouse small intestine. *Proc R Soc Lond B Biol Sci* (1990) 241:13–8. doi:10.1098/rspb.1990.0059
- Potten CS, Loeffler M. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* (1990) 110:1001–20.
- Nishimura S, Wakabayashi N, Toyoda K, Kashima K, Mitsufuji S. Expression of Musashi-1 in human normal colon crypt cells. *Dig Dis Sci* (2003) 48:1523–9. doi:10.1023/A:1024763723240
- Huang EH, Hynes MJ, Zhang T, Ginestier C, Dontu G, Appelman H, et al. Aldehyde dehydrogenase 1 is a marker for normal and malignant human colonic stem cells (SC) and tracks SC overpopulation during colon tumorigenesis. *Cancer Res* (2009) 69:3382–9. doi:10.1158/0008-5472.CAN-08-4418
- Roberts SA, Hendry JH, Potten CS. Deduction of the clonogen content of intestinal crypts: a direct comparison of two-dose and multiple-dose methodologies. *Radiat Res* (1995) 141:303–8. doi:10.2307/3579007
- Brittan M, Wright NA. Stem cell in gastrointestinal structure and neoplastic development. *Gut* (2004) 53:899–910. doi:10.1136/gut.2003.025478
- Boman BM, Huang E. Human colon cancer stem cells: a new paradigm in gastrointestinal oncology. *J Clin Oncol* (2008) 26:2828–38. doi:10.1200/JCO.2008.17.6941
- Deschner EE, Lewis CM, Lipkin M. In vitro study of human epithelial cells. I. Atypical zone of H3-thymidine incorporation in mucosa of multiple polyposis. *J Clin Invest* (1963) 42:1922–8. doi:10.1172/JCI104878
- Bleiberg H, Mainguet P, Galand P. Cell renewal in familial polyposis. Comparison between polyps and adjacent healthy mucosa. *Gastroenterology* (1972) 63:240–5.
- Iwana T, Utsunomiya J, Sasaki J. Epithelial cell kinetics in the crypts of familial polyposis of the colon. *Jpn J Surg* (1977) 7:230–4. doi:10.1007/BF02469355
- Boman BM, Fields JZ, Cavanaugh KL, Gujett A, Runquist OA. How dysregulated colonic crypt dynamics cause stem cell overpopulation and initiate colon cancer. *Cancer Res* (2008) 68:3304–13. doi:10.1158/0008-5472.CAN-07-2061
- Buczacki SJ, Zecchini HI, Nicholson AM, Russell R, Vermeulen L, Kemp R, et al. Intestinal label-retaining cells are secretory precursors expressing *Lgr5*. *Nature* (2013) 495:65–9. doi:10.1038/nature11965
- Cheshier SH, Morrison SJ, Liao X, Weissman IL. In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. *Proc Natl Acad Sci U S A* (1999) 96:3120–5. doi:10.1073/pnas.96.6.3120
- Randall TD, Weissman IL. Characterization of a population of cells in the bone marrow that phenotypically mimic hematopoietic stem cells: resting stem cells or mystery population? *Stem Cells* (1998) 16:38–48. doi:10.1002/stem.160038
- Boman BM, Wicha MS. Cancer stem cells: a step toward the cure. *J Clin Oncol* (2008) 26:2795–9. doi:10.1200/JCO.2008.17.7436
- 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) 445:111–5. doi:10.1038/nature05384

29. 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. doi:10.1038/nature05372
30. Dalerba P, Dylla SJ, Park IK, Liu R, Wang X, Cho RW, et al. Phenotypic characterization of human colorectal cancer stem cells. *Proc Natl Acad Sci U S A* (2007) **104**:10158–63. doi:10.1073/pnas.0703478104
31. Levin TG, Powell AE, Davies PS, Silk AD, Dismuke AD, Anderson EC, et al. Characterization of the intestinal cancer stem cell marker, CD166/ALCAM, in the human and mouse gastrointestinal tract. *Gastroenterology* (2010) **139**:2072–82. doi:10.1053/j.gastro.2010.08.053
32. Carpentino JE, Hynes MJ, Appelmann HD, Zheng T, Steindler DA, Scott EW, et al. Aldehyde dehydrogenase expressing colon stem cells contribute to tumorigenesis in the transition from colitis to cancer. *Cancer Res* (2009) **69**:8208–15. doi:10.1158/0008-5472.CAN-09-1132
33. Lin L, Liu A, Peng Z, Lin HJ, Li PK, Li C, et al. STAT3 is necessary for proliferation and survival in colon cancer-initiating cells. *Cancer Res* (2011) **71**:7226–37. doi:10.1158/0008-5472.CAN-10-4660
34. Bu P, Chen KY, Chen JH, Wang L, Walters J, Shin YJ, et al. A microRNA miR-34a-regulated bimodal switch targets notch in colon cancer stem cells. *Cell Stem Cell* (2013) **12**:602–15. doi:10.1016/j.stem.2013.03.002
35. Ro S, Rannala B. Methylation patterns and mathematical models reveal dynamics of stem cell turnover in the human colon. *Proc Natl Acad Sci U S A* (2001) **98**:10519–21. doi:10.1073/pnas.201405498
36. Molofsky AV, Pardo R, Morrison SJ. Diverse mechanisms regulate stem cell self-renewal. *Curr Opin Cell Biol* (2004) **16**:700–7. doi:10.1016/j.ccb.2004.09.004
37. Boman BM, Wicha M, Fields JZ, Runquist O. Symmetric division of cancer stem cells – a key mechanism in tumor growth that should be targeted in future therapeutic approaches. *Clin Pharmacol Ther* (2007) **81**:893–8. doi:10.1038/sj.clpt.6100202
38. Booth C, Potten CS. Gut instincts: thoughts on intestinal epithelial stem cells. *J Clin Invest* (2000) **105**:1493–9. doi:10.1172/JCI10229
39. Bjerknes M, Cheng H. Intestinal epithelial stem cells and progenitors. *Methods Enzymol* (2006) **419**:337–83. doi:10.1016/S0076-6879(06)19014-X
40. Bjerknes M, Cheng H. Neurogenin 3 and the enteroendocrine cell lineage in the adult mouse small intestinal epithelium. *Dev Biol* (2006) **300**:722–35. doi:10.1016/j.ydbio.2006.07.040
41. Van Landeghem L, Santoro MA, Krebs AE, Mah AT, Dehmer JJ, Gracz AD, et al. Activation of two distinct Sox9-EGFP-expressing intestinal stem cell populations during crypt regeneration after irradiation. *Am J Physiol Gastrointest Liver Physiol* (2012) **302**:G1111–32. doi:10.1152/ajpgi.00519.2011
42. Carlone DL, Breault DT. Tales from the Crypt: the expanding role of slow cycling intestinal stem cells. *Cell Stem Cell* (2012) **10**:2–4. doi:10.1016/j.stem.2011.12.012
43. Takeda N, Jain R, LeBoeuf MR, Wang Q, Lu MM, Epstein JA. Interconversion between intestinal stem cell populations in distinct niches. *Science* (2011) **334**:1420–4. doi:10.1126/science.1213214
44. 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. doi:10.1126/science.1145720
45. Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* (2012) **487**:330–7. doi:10.1038/nature11252
46. Powell SM, Zilz N, Beazer-Barclay Y, Bryan TM, Hamilton SR, Thibodeau SN, et al. APC mutations occur early during colorectal carcinogenesis. *Nature* (1992) **359**:235–7. doi:10.1038/359235a0
47. Levy DB, Smith KJ, Beazer-Barclay Y, Hamilton SR, Vogelstein B, Kinzler KW. Inactivation of both APC alleles in human and mouse tumors. *Cancer Res* (1994) **54**:5953–8.
48. Fearhead NS, Wilding JL, Bodmer WF. Genetics of colorectal cancer: hereditary aspects and overview of colorectal tumorigenesis. *Br Med Bull* (2002) **64**:27–43. doi:10.1093/bmb/64.1.27
49. Albuquerque C, Breukel C, van der Luijt R, Fidalgo P, Lage P, Slors FJM, et al. The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the β -catenin signaling cascade. *Hum Mol Genet* (2002) **11**:1549–60. doi:10.1093/hmg/11.13.1549
50. Potten CS, Kellett M, Rew DA, Roberts SA. Proliferation in human gastrointestinal epithelium using bromodeoxyuridine in vivo: data for different sites, proximity to a tumour, and polyposis coli. *Gut* (1992) **33**:524–9. doi:10.1136/gut.33.4.524
51. Lightdale C, Lipkin M, Deschner E. In vivo measurements in familial polyposis: kinetics and location of proliferating cells in colonic adenomas. *Cancer Res* (1982) **42**:4280–3.
52. Maskens AP. Histogenesis of adenomatous polyps in the human large intestine. *Gastroenterology* (1979) **77**:1245–51.
53. Cairnie AB, Bentley RE. Cell proliferation studies in the intestinal epithelium of the rat. Hyperplasia during lactation. *Exp Cell Res* (1967) **46**:428–40. doi:10.1016/0014-4827(67)90079-1
54. Totafurno J, Bjerknes M, Cheng H. Variation in crypt size and its influence on the analysis of epithelial cell proliferation in the intestinal crypt. *Biophys J* (1988) **54**:845–58. doi:10.1016/S0006-3495(88)83021-2
55. Totafurno J, Bjerknes M, Cheng H. Estimation of the extent of the cutoff region from the spatial distribution of labelling and mitotic indices of intestinal crypts of a fixed length. *Bull Math Biol* (1993) **55**:37–73. doi:10.1007/BF02460294
56. Lipkin M, Blattner WE, Fraumeni JF Jr, Lynch HT, Deschner E, Winaver S. Tritiated thymidine (ϕ , ϕ ,h) labeling distribution as a marker for hereditary predisposition to colon cancer. *Cancer Res* (1983) **43**:1899–904.
57. Lipkin M, Blattner WA, Gardner EJ, Burt RW, Lynch H, Deschner E, et al. Classification and risk assessment of individuals with familial polyposis, Gardner's syndrome, and familial non-polyposis colon cancer from [3H]thymidine labeling patterns in colonic epithelial cells. *Cancer Res* (1984) **44**:4201–7.
58. Zhang T, Fields JZ, Opendaker L, Otevrete T, Masuda E, Palazzo JP, et al. Survivin-induced Aurora-B kinase activation – a mechanism by which APC mutations contribute to increased mitoses during colon cancer development. *Am J Pathol* (2010) **177**:2816–26. doi:10.2353/ajpath.2010.100047
59. Mills SJ, Shepherd NA, Hall PA, Hastings A, Mathers JC, Gunn A. Proliferative compartment deregulation in the non-neoplastic colonic epithelium of familial adenomatous polyposis. *Gut* (1995) **36**:391–4. doi:10.1136/gut.36.3.391
60. Wasan HS, Park HS, Liu KC, Mandir NK, Winnett A, Sasieni P, et al. APC in the regulation of intestinal crypt fission. *J Pathol* (1998) **185**:246–55. doi:10.1002/(SICI)1096-9896(199807)185:3<246::AID-PATH90>3.3.CO;2-#
61. Mills SJ, Mathers JC, Chapman PD, Burn J, Gunn A. Colonic crypt cell proliferation state assessed by whole crypt microdissection in sporadic neoplasia and familial adenomatous polyposis. *Gut* (2001) **48**:41–6. doi:10.1136/gut.48.1.41
62. Boman BM, Fields JZ, Bonham-Carter O, Runquist OA. Computer modeling implicates stem cell overproduction in colon cancer initiation. *Cancer Res* (2001) **61**:8408–11.
63. Boman BM, Kopelovich L, Siracusa LD, Zhang T, Henderson K, Cofer Z, et al. A Tcf4-GFP reporter mouse model for monitoring effects of Apc mutations during intestinal tumorigenesis. *Mol Carcinog* (2009) **48**:821–31. doi:10.1002/mc.20526
64. Boman BM, Walters R, Fields JZ, Kovatich AJ, Zhang T, Isenberg GA, et al. Colonic crypt changes during adenoma development in familial adenomatous polyposis: immunohistochemical evidence for expansion of the crypt base cell population. *Am J Pathol* (2004) **165**:1489–98. doi:10.1016/S0002-9440(10)63407-4
65. Kim KM, Calabrese P, Tavaré S, Shibata D. Enhanced stem cell survival in familial adenomatous polyposis. *Am J Pathol* (2004) **164**:1369–77. doi:10.1016/S0002-9440(10)63223-3
66. McClanahan T, Koseoglu S, Smith K, Grein J, Gustafson E, Black S, et al. Identification of overexpression of orphan G protein-coupled receptor GPR49 in human colon and ovarian primary tumors. *Cancer Biol Ther* (2006) **5**(4):419–26.
67. Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, Cozijnsen M, et al. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature* (2007) **449**:1003–7. doi:10.1038/nature06196

68. Korinek V, Barker N, Moerer P, van Donselaar E, Huls G, Peters PJ. Depletion of epithelial stem cell compartments in the small intestine of mice lacking Tcf4. *Nat Genet* (1998) **19**:379–83. doi:10.1038/1270
69. Kielman MF, Rindapaa M, Gaspar C, van Poppel N, Breukel C, van Leeuwen S, et al. Apc modulates embryonic stem-cell differentiation by controlling the dosage of β -catenin signaling. *Nat Genet* (2002) **32**:594–605. doi:10.1038/ng1045
70. van de Wetering M, Sancho E, Verweij C, de Lau W, Oving I, Hurlstone A, et al. The β -catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* (2002) **111**:241–50. doi:10.1016/S0092-8674(02)01014-0
71. Barker N, Ridgway RA, van Es JH, van de Wetering M, Begthel H, van den Born M, et al. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* (2009) **457**:608–12. doi:10.1038/nature07602
72. Midgley CA, White S, Howitt R, Save V, Dunlop MG, Hall PA, et al. APC expression in normal human tissues. *J Pathol* (1997) **181**:426–33. doi:10.1002/(SICI)1096-9896(199704)181:4<426::AID-PATH768>3.3.CO;2-K
73. Miyashiro I, Senda T, Matsumine A, Baeg G, Kuroda T, Shimano T, et al. Subcellular localization of the APC protein: immunoelectron microscopic study of the association of the APC protein with catenin. *Oncogene* (1995) **11**:89–96.
74. Smith KJ, Johnson KA, Bryan TM, Hill DE, Markowitz S, Willson JKV, et al. The APC gene product in normal and tumor cells. *Proc Natl Acad Sci U S A* (1993) **90**:2846–50. doi:10.1073/pnas.90.7.2846
75. Umar S, Wang Y, Sellin JH. Epithelial proliferation induces novel changes in APC expression. *Oncogene* (2005) **24**:6709–18. doi:10.1038/sj.onc.1208820
76. Reinacher-Schick A, Gumbiner BM. Apical membrane localization of the APC tumor suppressor protein and subcellular distribution of the β -catenin destruction complex in polarized epithelial cells. *J Cell Biol* (2001) **152**:491–502. doi:10.1083/jcb.152.3.491
77. Näthke IS, Adams CL, Polakis P, Sellin JH, Nelson WJ. The adenomatous polyposis coli tumor suppressor protein localizes to plasma membrane sites involved in active cell migration. *J Cell Biol* (1996) **134**:165–79. doi:10.1083/jcb.134.1.165
78. Senda T, Miyashiro I, Matsumine A, Baeg GH, Monden T, Kobayashi S, et al. The tumor suppressor protein APC colocalizes with β -catenin in the colon epithelial cells. *Biochem Biophys Res Commun* (1996) **223**:329–34. doi:10.1006/bbrc.1996.0894
79. Maltzman T, Whittington J, Driggers L, Stephens J, Ahnen D. AOM-induced mouse colon tumors do not express full-length APC protein. *Carcinogenesis* (1996) **18**:2435–9. doi:10.1093/carcin/18.12.2435
80. Wang Y, Azuma Y, Friedman DB, Coffey RJ, Neufeld KL. Novel association of APC with intermediate filaments identified using a new versatile APC antibody. *BMC Cell Biol* (2009) **10**:75. doi:10.1186/1471-2121-10-75
81. Holcombe RF, Marsh JL, Waterman ML, Lin F, Milovanovic T, Truong T. Expression of Wnt ligands and Frizzled receptors in colonic mucosa and in colon carcinoma. *Mol Pathol* (2002) **55**:220–6. doi:10.1136/mp.55.4.220
82. Gregorieff A, Pinto D, Begthel H, Destree O, Kielman M, Clevers H. Expression pattern of Wnt signaling components in the adult intestine. *Gastroenterology* (2005) **129**:626–38. doi:10.1016/j.gastro.2005.06.007
83. Kongkanunt R, Bubbs VJ, Sansom OJ, Wyllie AH, Harrison DJ, Clarke AR. Dysregulated expression of β -catenin marks early neoplastic change in Apc mutant mice, but not all lesions arising in Msh2 deficient mice. *Oncogene* (1999) **18**:7219–25. doi:10.1038/sj.onc.1203181
84. Sellin JH, Umar S, Xiao J, Morris AP. Increased β -catenin expression and nuclear translocation accompany cellular hyperproliferation in vivo. *Cancer Res* (2001) **61**:2899–906.
85. Fodde R, Smits R, Clevers H. APC, signal transduction and genetic instability in colorectal cancer. *Nat Rev Cancer* (2001) **1**:55–67. doi:10.1038/35094067
86. Gregorieff A, Clevers H. Wnt signaling in the intestinal epithelium: from endoderm to cancer. *Genes Dev* (2005) **19**:877–90.
87. Fevr T, Robine S, Louvard D, Huelsken J. Wnt/ β -catenin is essential for intestinal homeostasis and maintenance of intestinal stem cells. *Mol Cell Biol* (2007) **27**:7551–9. doi:10.1128/MCB.01034-07
88. Fagman H, Larsson F, Arvidsson Y, Meuller J, Nordling M, Martinsson T, et al. Nuclear accumulation of full-length and truncated adenomatous polyposis coli protein in tumor cells depends on proliferation. *Oncogene* (2003) **22**:6013–22. doi:10.1038/sj.onc.1206731
89. Deng G, Song GA, Pong E, Sleisenger M, Kim YS. Promoter methylation inhibits APC gene expression by causing changes in chromatin conformation and interfering with the binding of transcription factor CCAAT-binding factor. *Cancer Res* (2004) **64**:2692–8. doi:10.1158/0008-5472.CAN-03-3000
90. Virmani AK, Rath A, Sathyanarayana UG, Padar A, Huang CX, Cunningham HT, et al. Aberrant methylation of the adenomatous polyposis coli (APC) gene promoter 1A in breast and lung carcinomas. *Clin Cancer Res* (2001) **7**:1998–2004.
91. Jaiswal AS, Narayan S. Protein synthesis and transcriptional inhibitors control N-methyl-N'-nitro-N-nitrosoguanidine-induced levels of APC mRNA in a p53-dependent manner. *Int J Oncol* (1998) **13**:733–40.
92. Jaiswal AS, Narayan S. Upstream stimulating factor-1 (USF1) and USF2 bind to and activate the promoter of the adenomatous polyposis coli (APC) tumor suppressor gene. *J Cell Biochem* (2001) **81**:262–77. doi:10.1002/1097-4644(20010501)81:2<262::AID-JCB1041>3.0.CO;2-R
93. Jaiswal AS, Narayan S. p53-dependent transcriptional regulation of the APC promoter in colon cancer cells treated with DNA alkylating agents. *J Biol Chem* (2001) **276**:18193–9. doi:10.1074/jbc.M101298200
94. Jaiswal AS, Balusuy R, Narayan S. 7,12-Dimethylbenzanthracene-dependent transcriptional regulation of adenomatous polyposis coli (APC) gene expression in normal breast epithelial cells is mediated by GC-box binding protein Sp3. *Carcinogenesis* (2006) **27**:252–61. doi:10.1093/carcin/bgi225
95. Narayan S, Jaiswal AS. Activation of adenomatous polyposis coli (APC) gene expression by the DNA-alkylating agent N-methyl-N'-nitro-N-nitrosoguanidine requires p53. *J Biol Chem* (1997) **272**:30619–22. doi:10.1074/jbc.272.49.30619
96. Willert K, Jones KA. Wnt signaling: is the party in the nucleus? *Genes Dev* (2006) **20**:1394–404. doi:10.1101/gad.1424006
97. Hamilton PW, Bartels PH, Wilson RH, Sloan JM. Nuclear texture measurements in normal colorectal glands. *Anal Quant Cytol Histol* (1995) **17**:397–405.
98. Caldwell CM, Kaplan KB. The role of APC in mitosis and in chromosome instability. In: Näthke IS, editor. *APC Proteins*. New York: Springer Science+Business Media, LLC (2009). p. 51–64.
99. Bahmanyar S, Nelson WJ, Barth AIM. Role of APC and its binding partners in regulating microtubules in mitosis. In: Näthke IS, McCartney BM, editors. *APC Proteins*. New York, NY: Springer Science+Business Media, LLC (2009). p. 61–74.
100. Zhang T, Otevrel T, Gao ZQ, Gao ZP, Ehrlich SM, Fields JZ, et al. Evidence that APC regulates survivin expression: a possible mechanism contributing to the stem cell origin of colon cancer. *Cancer Res* (2001) **62**:8664–7.
101. Li F, Ambrosini G, Chu EY, Plescia J, Tognin S, Marchisio PC, et al. Control of apoptosis and mitotic spindle checkpoint by survivin. *Nature* (1998) **396**:580–4. doi:10.1038/25141
102. Altieri DC. The case for survivin as a regulator of microtubule dynamics and cell-death decisions. *Curr Opin Cell Biol* (2006) **18**:609–15. doi:10.1016/j.ceb.2006.08.015
103. Rosa J, Canovas P, Islam A, Altieri DC, Dosssey SJ. Survivin modulates microtubule dynamics and nucleation through the cell cycle. *Mol Biol Cell* (2006) **17**:1483–93. doi:10.1091/mbc.E05-08-0723
104. Kim PJ, Plescia J, Clevers H, Fearon ER, Altieri DC. Survivin and molecular pathogenesis of colorectal cancer. *Lancet* (2003) **362**:205–9. doi:10.1016/S0140-6736(03)13910-4
105. Boman BM, Zhang T, Fields JZ. Correspondence: evidence that APC regulates survivin expression: a possible mechanism contributing to the stem cell origin of colon cancer. *Cancer Res* (2004) **64**:776–9.
106. Ma H, Nguyen C, Lee KS, Kahn M. Differential roles for the coactivators CBP and p300 on TCF/ β -catenin-mediated survivin gene expression. *Oncogene* (2005) **24**:3619–31. doi:10.1038/sj.onc.1208433

107. Munemitsu S, Albert I, Souza B, Rubinfeld B, Polakis P. Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. *Proc Natl Acad Sci U S A* (1995) **92**:3046–50. doi:10.1073/pnas.92.7.3046
108. Morrison EE. The APC-EB1 interaction. In: Nathke IS, McCartney BM, editors. *APC Proteins*. New York, NY: Springer Science+Business Media, LLC (2009). p. 41–50.
109. Hansen CA, Miller JR. Non-traditional roles for the adenomatous polyposis coli (APC) tumor suppressor protein. *Gene* (2005) **361**:1–12. doi:10.1016/j.gene.2005.07.024
110. Goto H, Yasui Y, Kawajiri A, Nigg EA, Terada Y, Tatsuka M, et al. Aurora-B regulates the cleavage furrow-specific vimentin phosphorylation in the cyto-kinetic process. *J Biol Chem* (2003) **278**:8526–30. doi:10.1074/jbc.M210892200
111. Kawajiri A, Yasui Y, Goto H, Tatsuka M, Takahashi M, Nagata K, et al. Functional significance of the specific sites phosphorylated in desmin at cleavage furrow: aurora-B may phosphorylate and regulate Type III intermediate filaments during cytokinesis coordinatedly with Rho-kinase. *Mol Biol Cell* (2003) **14**:1489–500. doi:10.1091/mbc.E02-09-0612
112. Sun L, Gao J, Dong X, Liu M, Li D, Shi X, et al. EBI promotes aurora-B kinase activity through blocking its inactivation by protein phosphatase 2A. *Proc Natl Acad Sci U S A* (2008) **105**:7153–8. doi:10.1073/pnas.0710018105
113. Mao Y. FORMIN a link between kinetochores and microtubules. *Trends Cell Biol* (2011) **21**:625–9. doi:10.1016/j.tcb.2011.08.005
114. Cheng L, Zhang J, Ahmad S, Rozier L, Yu H, Dang H, et al. Aurora B regulates formin mDia3 in achieving metaphase chromosome alignment. *Cell* (2011) **145**:342–52. doi:10.1016/j.devcel.2011.01.008
115. Sieber OM, Heinemann K, Gorman P, Lamlum H, Crabtree M, Simpson CA, et al. Analysis of chromosomal instability in human colorectal adenomas with two mutational hits at APC. *Proc Natl Acad Sci U S A* (2002) **99**:16910–5. doi:10.1073/pnas.012679099
116. Crabtree M, Sieber OM, Lip-ton L, Hodgson SV, Lamlum H, Thomas HJW, et al. Refining the relation between ‘first hits’ and ‘second hits’ at the APC locus: the ‘loose fit’ model and evidence for differences in somatic mutation spectra among patients. *Oncogene* (2003) **22**:4257–65. doi:10.1038/sj.onc.1206471
117. Lam-lum H, Ilyas M, Rowan A, Clark S, Johnson V, Bell J, et al. The type of somatic mutation at APC in familial adenomatous polyposis is determined by the site of the germline mutation: a new facet to Knudson’s ‘two-hit’ hypothesis. *Nat Med* (1999) **5**:1071–5. doi:10.1038/12511
118. Rowan AJ, Lam-lum H, Ilyas M, Wheeler J, Straub J, Papadopoulos A, et al. APC mutations in sporadic colorectal tumors: a mutational “hotspot” and interdependence of the “two hits”. *Proc Natl Acad Sci U S A* (2000) **97**:3352–7. doi:10.1073/pnas.97.7.3352
119. Fearhead NS, Britton MP, Bodmer WF. The ABC of APC. *Hum Mol Genet* (2001) **10**:721–33. doi:10.1093/hmg/10.7.721
120. Christie M, Jorissen RN, Mouradov D, Sakthianandeswaren A, Li S, Day F, et al. Different APC genotypes in proximal and distal sporadic colorectal cancers suggest distinct WNT/beta-catenin signaling thresholds for tumorigenesis. *Oncogene* (2012) **32**:4675–82. doi:10.1038/onc.2012.486
121. Quyn AJ, Appleton PL, Carey FA, Steele RJ, Barker N, Clevers H, et al. Spindle orientation bias in gut epithelial stem cell compartments is lost in precancerous tissue. *Cell Stem Cell* (2010) **6**:175–81. doi:10.1016/j.stem.2009.12.007
122. Sherley JL. Asymmetric self-renewal: the mark of the adult stem cell. In: Habib NA, Gordon MY, Levicar N, Jiao L, Thomas-Black G, editors. *Stem Cell Repair and Regeneration*. London: Imperial College Press (2005). p. 21–8.
123. Doupe DP, Jones PH. Cycling progenitors maintain epithelia while diverse cell types contribute to repair. *Bioessays* (2013) **35**:443–51. doi:10.1002/bies.201200166
124. Yatabe Y, Taware S, Shibata D. Investigating stem cells in human colon by using methylation patterns. *Proc Natl Acad Sci U S A* (2001) **98**:10839–44. doi:10.1073/pnas.191225998
125. Snippert HJ, van der Flier LG, Sato T, van Es JH, van den Born M, Kroon-Veenboer C, et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* (2010) **143**:134–44. doi:10.1016/j.cell.2010.09.016
126. Lopez-Garcia C, Klein AM, Simons BD, Winton DJ. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* (2010) **330**:822–5. doi:10.1126/science.1196236
127. Schepers AG, Vries R, van den Born M, van de Wetering M, Clevers H. Lgr5 intestinal stem cells have high telomerase activity and randomly segregate their chromosomes. *EMBO J* (2011) **30**:1104–9. doi:10.1038/emboj.2011.26
128. Klein AM, Simons BD. Universal patterns of stem cell fate in cycling adult tissues. *Development* (2011) **138**:3103–11. doi:10.1242/dev.060103
129. Cairns J. Mutation selection and the natural history of cancer. *Nature* (1975) **255**:197–200. doi:10.1038/255197a0
130. Potten CS, Owen G, Booth D. Intestinal stem cells protect their genome by selective segregation of template DNA strands. *J Cell Sci* (2002) **115**:2381–8.
131. Merok JR, Lansita JA, Tunstead JR, Sherley JL. Cosegregation of chromosomes containing immortal DNA strands in cells that cycle with asymmetric stem cell kinetics. *Cancer Res* (2002) **62**:6791–5.
132. Falconer E, Chavez EA, Henderson A, Poon SS, McKinney S, Brown L, et al. Identification of sister chromatids by DNA template strand sequences. *Nature* (2010) **463**:93–7. doi:10.1038/nature08644
133. Yan KS, Chia LA, Xingnan L, Ootani A, Su J, Lee JY, et al. The intestinal stem cell markers Bmi1 and Lgr5 identify two functionally distinct populations. *Proc Natl Acad Sci U S A* (2012) **109**:466–71. doi:10.1073/pnas.1118857109
134. Barker N, van Es JH, Kakes V, Kasper M, Snippert H, Toftgard R, et al. Very long-term self-renewal of small intestine, colon and hair follicles from cycling Lgr5+ve stem cells. *Cold Spring Harb Symp Quant Biol* (2010) **73**:351–6. doi:10.1101/sqb.2008.72.003
135. Powell AE, Wang Y, Li Y, Poulin EJ, Means AL, Washington MK, et al. The pan-ErbB negative regulator Lrig1 is an intestinal stem cell marker that functions as a tumor suppressor. *Cell* (2012) **149**:146–58. doi:10.1016/j.cell.2012.02.042
136. Montgomery RK, Carlone DL, Richmond CA, Farilla LL, Kranendonk EG, Henderson DE, et al. Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. *Proc Natl Acad Sci U S A* (2011) **108**:179–84. doi:10.1073/pnas.1013004108
137. Tian H, Biehs B, Warming S, Leong KG, Rangell L, Klein OD, et al. A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. *Nature* (2011) **478**:255–60. doi:10.1038/nature10408
138. Schaniel C, Moore KA. Genetic models to study quiescent stem cells and their niches. *Ann NY Acad Sci* (2009) **1176**:26–35. doi:10.1111/j.1749-6632.2009.04608.x
139. Wilson A, Laurenti E, Oser G, van der Wath RC, Blanco-Bose W, Jaworski M, et al. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* (2008) **135**:1118–29. doi:10.1016/j.cell.2008.10.048
140. Sirin O, Lukov GL, Mao R, Conneely OM, Goodell MA. The orphan nuclear receptor Nurrl restricts the proliferation of haematopoietic stem cells. *Nat Cell Biol* (2010) **12**:1213–9. doi:10.1038/ncb2125
141. Winton DA. miR-34a sets the sweet spot for notch in colorectal cancer stem cells. *Cell Stem Cell* (2013) **12**:499–501. doi:10.1016/j.stem.2013.04.018
142. Snippert HJ, Clevers H. Tracking adult stem cells. *EMBO Rep* (2011) **12**:113–22. doi:10.1038/embor.2010.216
143. Bellis J, Duluc I, Romagnolo B, Perret C, Faux MC, Dujardin D, et al. The tumor suppressor Apc controls planar cell polarities central to gut homeostasis. *J Cell Biol* (2012) **198**:331–41. doi:10.1083/jcb.201204086
144. Kobayashi S, Yamada-Okabe H, Suzuki M, Natori O, Kato A, Matsubara K, et al. LGR5-positive colon cancer stem cells interconvert with drug-resistant LGR5-negative cells and are capable of tumor reconstitution. *Stem Cells* (2012) **30**:2631–44. doi:10.1002/stem.1257
145. Glauche I, Moore K, Thielecke L, Horn K, Loeffler M, Roeder I. Stem cell proliferation and quiescence – two sides of the same coin. *PLoS Comput Biol* (2009) **5**:e1000447. doi:10.1371/journal.pcbi.1000447

146. Bjerknes M. A test of the stochastic theory of stem cell differentiation. *Biophys J* (1986) **49**:1223–7. doi:10.1016/S0006-3495(86)83751-1
147. Totafurno J, Bjerknes M, Cheng H. The crypt cycle: crypt and villous production in the adult intestinal epithelium. *Biophys J* (1987) **52**:279–94. doi:10.1016/S0006-3495(87)83215-0
148. Bjerknes M. Simple stochastic theory of stem cell differentiation is not simultaneously consistent with crypt extinction probability and the expansion of mutated clones. *J Theor Biol* (1994) **168**:349–65. doi:10.1006/jtbi.1994.1115
149. Bjerknes M. The crypt cycle and the asymptotic dynamics of the proportion of differently sized mutant crypt clones in the mouse intestine. *Proc Biol Sci* (1995) **260**:1–6. doi:10.1098/rspb.1995.0051
150. Cheng H, Bjerknes M, Amar J, Gardiner G. Crypt production in normal and diseased human colonic epithelium. *Anat Rec* (1986) **216**:44–8. doi:10.1002/ar.1092160108
151. Araki K, Ogata T, Kobayashi M, Yatani R. A morphological study on the histogenesis of human colorectal hyperplastic polyps. *Gastroenterology* (1995) **109**:1468–74. doi:10.1016/0016-5085(95)90632-0
152. Cummins AG, Catto-Smith AG, Cameron DJ, Couper RT, Davidson GP, Day AS, et al. Crypt fission peaks early during infancy and crypt hyperplasia broadly peaks during infancy and childhood in the small intestine of humans. *J Pediatr Gastroenterol Nutr* (2008) **47**:153–7. doi:10.1097/MPG.0b013e3181604d27
153. Bjerknes M, Cheng H, Hay K, Gallinger S. APC mutation and the crypt cycle in murine and human intestine. *Am J Pathol* (1997) **150**:833–9.
154. Park HS, Goodlad RA, Ahnen DJ, Winnett A, Sasieni P, Lee CY, et al. Effects of epidermal growth factor and dimethylhydrazine on crypt size, cell proliferation, and crypt fission in the rat colon. Cell proliferation and crypt fission are controlled independently. *Am J Pathol* (1997) **151**:843–52.
155. Brittan M, Wright NA. Gastrointestinal stem cells. *J Pathol* (2002) **197**:492–509. doi:10.1002/path.1155
156. Wong WM, Mandir N, Goodlad RA, Wong BC, Garcia SB, Lam SK, et al. Histogenesis of human colorectal adenomas and hyperplastic polyps: the role of cell proliferation and crypt fission. *Gut* (2002) **50**:212–7. doi:10.1136/gut.50.2.212
157. McDonald SA, Preston SL, Greaves LC, Leedham SJ, Lovell MA, Jankowski JA, et al. Clonal expansion in the human gut: mitochondrial DNA mutations show us the way. *Cell Cycle* (2006) **5**:808–11. doi:10.4161/cc.5.8.2641
158. Leedham SJ, Thliveris AT, Halberg RB, Newton MA, Wright NA. Gastrointestinal stem cells and cancer: bridging the molecular gap. *Stem Cell Rev* (2005) **1**:233–41. doi:10.1385/SCR.1:3:233
159. Nakamura S, Kino I. Morphogenesis of minute adenomas in familial polyposis coli. *J Natl Cancer Inst* (1984) **73**:41–9.
160. Camac KS, Thompson FM, Cummins AG. Activation of beta-catenin in the stem cell region of crypts during growth of the small intestine in infant rats. *Dig Dis Sci* (2007) **52**:1242–6. doi:10.1007/s10620-006-9200-7
161. Alonso L, Fuchs E. The hair cycle. *J Cell Sci* (2006) **119**:391–3. doi:10.1242/jcs02793
162. Kishimoto J, Burgeson RE, Morgan BA. Wnt signaling maintains the hair-inducing activity of the dermal papilla. *Genes Dev* (2000) **14**:1181–5.
163. Maskens AP. Histogenesis of colon glands during post-natal growth. *Acta Anat (Basel)* (1978) **100**:17–26. doi:10.1159/000144877
164. Dehmer JJ, Garrison AP, Speck KE, Dekaney CM, Van Landeghem L, Sun X, et al. Expansion of intestinal epithelial stem cells during murine development. *PLoS One* (2011) **6**(11):e27070. doi:10.1371/journal.pone.0027070
165. Janssens N, Janicot M, Perera T. The WNT-dependent signaling pathways as targets in oncology drug discovery. *Invest New Drugs* (2006) **24**:263–80. doi:10.1007/s10637-005-5199-4
166. Watanabe K, Dai X. Winning WNT: race to WNT signaling inhibitors. *Proc Natl Acad Sci U S A* (2011) **108**:5929–30. doi:10.1073/pnas.1103102108

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 April 2013; accepted: 03 September 2013; published online: 07 November 2013.

Citation: Boman BM and Fields JZ (2013) An APC:WNT counter-current-like mechanism regulates cell division along the human colonic crypt axis: a mechanism that explains how APC mutations induce proliferative abnormalities that drive colon cancer development. *Front. Oncol.* **3**:244. doi:10.3389/fonc.2013.00244

This article was submitted to *Cancer Genetics*, a section of the journal *Frontiers in Oncology*.

Copyright © 2013 Boman and Fields. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Marrow hematopoietic stem cells revisited: they exist in a continuum and are not defined by standard purification approaches; then there are the microvesicles

Peter J. Quesenberry*, Laura Goldberg, Jason Aliotta and Mark Dooner

Department of Medicine, Division of Hematology/Oncology, Rhode Island Hospital, Providence, RI, USA

Edited by:

James L. Sherley, Boston Biomedical Research Institute, USA

Reviewed by:

Nadim Mahmud, University of Illinois at Chicago, USA

James L. Sherley, Boston Biomedical Research Institute, USA

*Correspondence:

Peter J. Quesenberry, Department of Medicine, Division of Hematology/Oncology, Rhode Island Hospital, 593 Eddy Street, Providence, RI 02903, USA
e-mail: pquesenberry@lifespan.org

Current concepts of hematopoiesis are encompassed in a hierarchical stem cell model. This developed initially from studies of colony-forming unit spleen and *in vitro* progenitors for different cell lineages, but then evolved into a comprehensive model of cells with different *in vivo* differentiative and proliferative potential. These cells were characterized and purified based largely on expression of a variety of lineage-specific and stem cell-specific surface epitopes. Monoclonal antibodies were bound to these epitopes and then used to physically and fluorescently separate different classes of these cells. The gold standard for the most primitive marrow stem cells was long-term multilineage repopulation and renewal in lethally irradiated mice. Progressive work seemed to have clonally defined a Lineage negative (Lin⁻), Sca-1⁺, c-kit⁺, CD150⁺ stem cell with great proliferative, differentiative, and renewal potential. This cell was stable and in the G0 phase of cell cycle. However, continued work in our laboratory indicated that the engraftment, differentiation, homing, and gene expression phenotype of the murine marrow stem cells continuously and reversibly changes with passage through cell cycle. Most recently, using cycle-defining supravital dyes and fluorescent-activated cell sorting and S-phase-specific tritiated thymidine suicide, we have established that the long-term repopulating hematopoietic stem cell is a rapidly proliferating, and thus a continually changing cell; as a corollary it cannot be purified or defined on a clonal single cell basis. Further *in vivo* studies employing injected and ingested 5-Bromodeoxyuridine (BrdU), showed that the G0 Lin⁻Sca-1⁺, c-kit⁺ Flt3⁻ cell was rapidly passing through cell cycle. These data are explained by considering the separative process: the proliferating stem cells are eliminated through the selective separations leaving non-representative dormant G0 stem cells. In other words, they throw out the real stem cells with the purification. This system, where the marrow stem cell continuously and reversibly changes with obligate cell cycle transit, is further complicated by the consideration of the impact of tissue microvesicles on the cell phenotypes. Tissue microvesicles have been found to alter the phenotype of marrow cells, possibly explaining the observations of "stem cell plasticity." These alterations, short-term, are due to transfer of originator cell mRNA and as yet undefined transcription factors. Long-term phenotype change is due to transcriptional modulation; a stable epigenetic change. Thus, the stem cell system is characterized by continuous cycle and microvesicle-related change. The challenge of the future is to define the stem cell population.

Keywords: stem cell, cell cycle, stem cell purification, vesicles, circadian rhythm

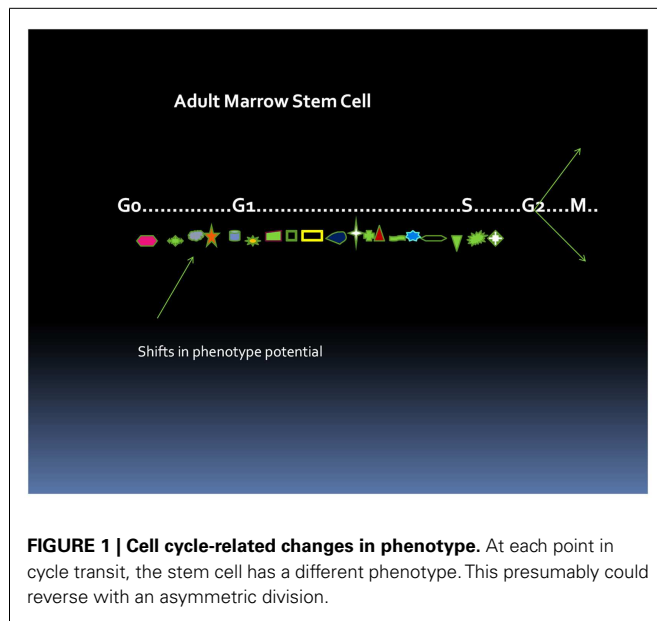
INTRODUCTION

NOTES ON CELL CYCLE AND CELL PHENOTYPE

The cell cycle status of a stem cell is a major determinant of cell phenotype and potential. A stem cell progressing through cell cycle will be continually changing its phenotype as to surface epitopes, RNA and DNA content, metabolic status, and overall phenotype and thus cannot be precisely characterized as a single entity (Figure 1).

The G0 state may be characterized by 2N DNA and low RNA levels with G1 showing increasing RNA levels and S showing increasing DNA levels. Mitosis then starts the whole sequence over

with the three conceptual outcomes: (1) a symmetric division in which each daughter cell retains its identity, (2) a symmetric division in which each daughter cell has differentiated, and (3) an asymmetric division in which one daughter cell has differentiated and in which the other has maintained its original identity. There is also the possibility of cell death of one or both daughters. In general, it is assumed that the final end result of divisions in the stem cell population should maintain the stem and differentiated populations on a steady state basis. An excess of symmetric divisions of stem cells would result in leukemia and an excess of differentiated end results would result in exhaustion and aplastic anemia. These

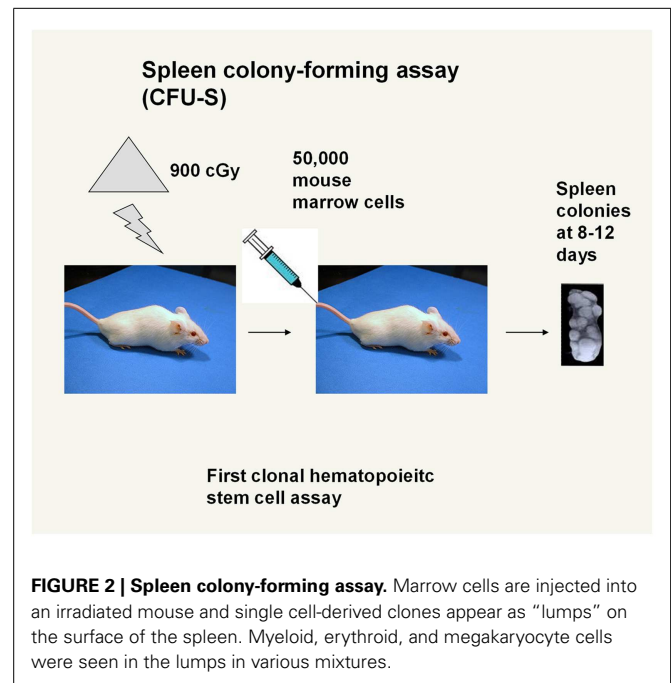


considerations are paramount to understanding current marrow stem cell biology.

THE FIRST CLONAL STEM CELL – THE COLONY-FORMING UNIT SPLEEN

The colony-forming unit spleen (CFU-S), as reported by Till and McCulloch (1), was the first description of a clonal stem cell unit. They received the Lasker award in 2005 for this seminal work on stem cells. While there has been a long period when this assay has been regarded as not defining the true stem cell, our own work would suggest that it indeed is a very good stem cell assay. Many of the original insights on the CFU-S would appear to be valid, in light of current work on adult marrow stem cells. The assay itself involves injecting murine marrow cells intravenously into lethally irradiated mice and then counting lumps (or clones) on the spleen which were stained with Bouin's fixative (Figure 2).

The cellular makeup of the colonies varied depending upon the location in the spleen, but these cells had the potential for differentiation into all myeloid cell classes and for self renewal. The CFU-S was characterized as a cell with an extensive capacity for differentiation and proliferation along with self renewal. It was shown that bumps on the spleen were clonal and that cells from a colony could form colonies in secondary irradiated hosts. Thus, the characteristics of the marrow stem cells were outlined as a cell which had extensive proliferative and differentiative potential into marrow myeloid cell types and which could self renew. An important feature of these early studies was observations of the heterogeneity of the formed colonies as to size, cell number, cell type, and location (2). Furthermore, it appeared that different cells might be monitored if the colonies were counted at 9, 12, or 14 days with more primitive cells being monitored by colonies, which arose at longer time intervals (3). Perhaps the most striking and important feature of the CFU-S was the total heterogeneity of self renewal from individual colonies. This would appear to be particularly relevant to our current concepts of the biology of adult marrow hematopoietic stem cells. In discussing how this lax

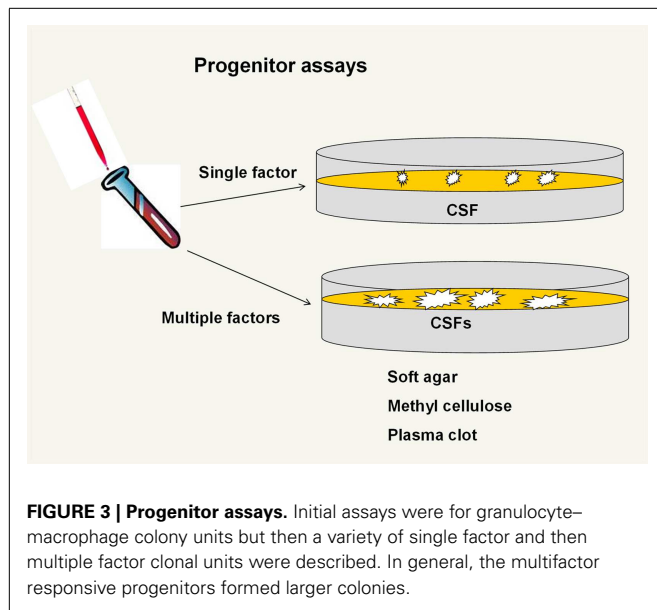


regulation (heterogeneity) could be reconciled with the orderly behavior of normal hematopoietic tissue, Till et al. (2) drew an analogy with radioactive atoms, “If one studies a large number of radioactive atoms, one sees a very regular pattern of decay following an exponential law. However, if one studies individual atoms, they are found to decay in an unpredictable fashion at random. It appears possible that our studies of the progeny of single cells display the random feature of hematopoietic function, while study of large populations of cells reveals the orderly behavior of the whole system. From this point of view, it is the population as a whole that is regulated rather than individual cells and it is suggested that control mechanisms act by varying the “birth” and “death” probabilities.” As we will develop below, these are very prescient comments, which apply to the current state of stem cell biology.

THE PROGENITOR ERA – IT FIT SO WELL

The next thrust of research in the stem cell field was the definition of progenitor classes of hematopoietic cells. Bradley and Metcalf (4) and Pluznik and Sachs (5) described the *in vitro* cloning in semisolid media of marrow cells that form granulocyte-macrophage colonies. As work here developed, the systems involved various semisolid matrices including soft agar, methyl cellulose, and plasma clot and various sources of “colony-stimulating factors” including mouse embryo-conditioned media, serum from endotoxin-treated mice, and cell feeder layers (Figure 3).

This work expanded as different investigators described cells giving rise to erythroid and megakaryocyte colonies (6) and then subsets of these lineage-specific colonies were described such that large colonies responding to multiple growth factors were termed burst-forming unit erythroid (7) and burst-forming unit megakaryocyte (8), while smaller colonies responding to one or a few cytokines were termed colony-forming unit erythroid or megakaryocyte. Relatively primitive cells giving rise to blast

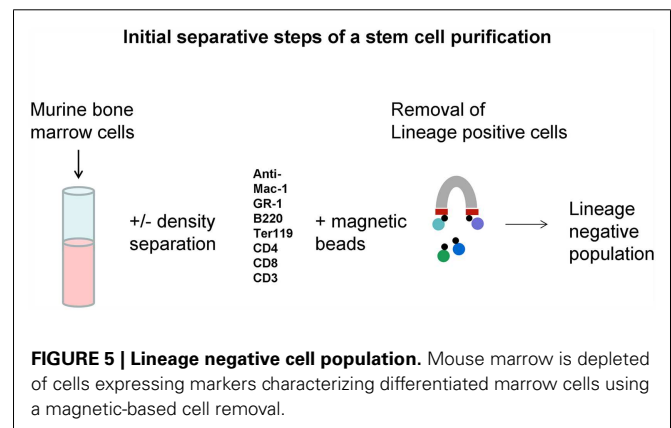
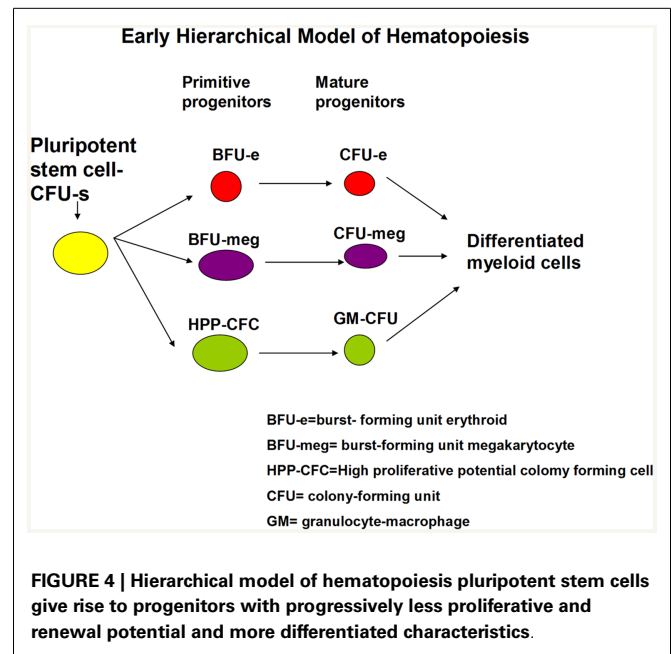


colonies (9) or high-proliferative potential colonies (10) were then defined and felt to possibly be surrogates for long-term repopulating marrow stem cells. Dr. Ogawa described a bewildering array of different colony types with from one to five lineages arising from single cells. Almost all possible combinations of differentiated cell colonies were seen (4). This gave rise to a hierarchical model with the multipotent CFU-S giving rise to multipotent progenitors (MPPs) with more limited potential which then, in turn, gave rise to bi or unipotent progenitors followed by recognizable differentiated myeloid cells. A simplified early hierarchical model is presented in **Figure 4**.

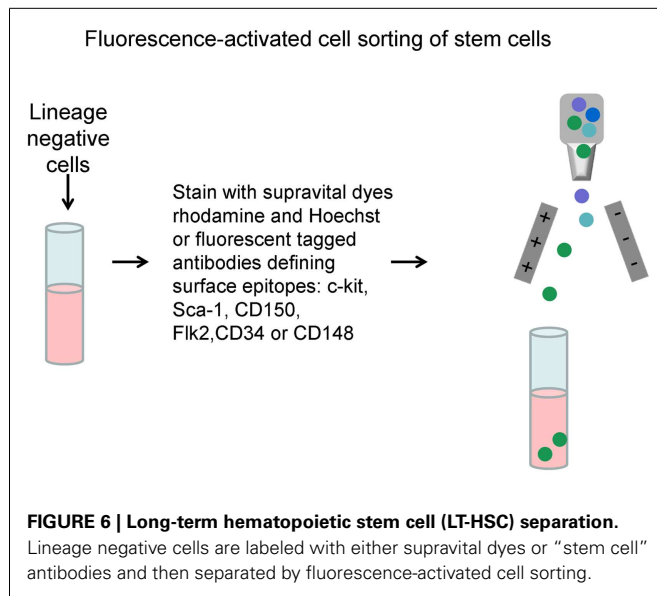
This suggested a very orderly system of hematopoiesis regulated by a series of cytokines or colony-stimulating factors with more primitive cells needing more factors to express their phenotype. Dr. Ogawa also published data showing that within one cell cycle transit from a blast colony-forming cell, totally different lineages could be pursued by the daughter cells (4). Thus, one daughter might give rise to a granulocyte–macrophage colony while the other daughter gave rise to an erythroid–megakaryocyte colony. The implications of these careful observations were generally ignored. These data were akin to throwing a bomb in the middle of any hierarchical model. As we will develop below, these data fit an alternative continuum model of hematopoiesis. With the definition of many progenitor cell classes, the emphasis of research turned to the precise clonal definition of the “true” long-term repopulating marrow stem cells and a full elucidation of the complex hematopoietic hierarchy.

THE PURIFICATORS

Early work suggested that cells with markers of differentiation had low to no long-term repopulating cells as defined by long-term multilineage repopulation in a lethally irradiated mouse. These studies were usually carried out using congenic mouse transplant models, the CD54.1 and CD45.2 strains being most often employed. They also frequently extended to secondary



repopulation in irradiated host to demonstrate “renewal.” Typically antibodies to differentiated cell markers with iron tags were incubated with marrow cells and positive “differentiated” cells removed by magnetic adherence. Then, this lineage negative population was incubated with antibodies to cell surface epitopes, the presence or absence of which, enriched for long-term repopulating cells (5). Many candidate stem cell markers were evaluated with positivity for c-kit, Sca-1, intermediate staining for Thy.1, and negativity for FLK2 (11–13) being initially defined markers and CD150 or Slam (14) and CD34 (15) also currently in vogue for definition of the stem cell. These studies showed a lack of tight correlation with the purified stem cells and CFU-S and led to the dismissal of CFU-S as a relevant stem cell; as we will develop below, this was probably a fundamental mistake. These studies also led to the evolved dogma that a stem cell could only be defined clonally. The general aspects of a stem cell separation are pictured in **Figures 5 and 6**.



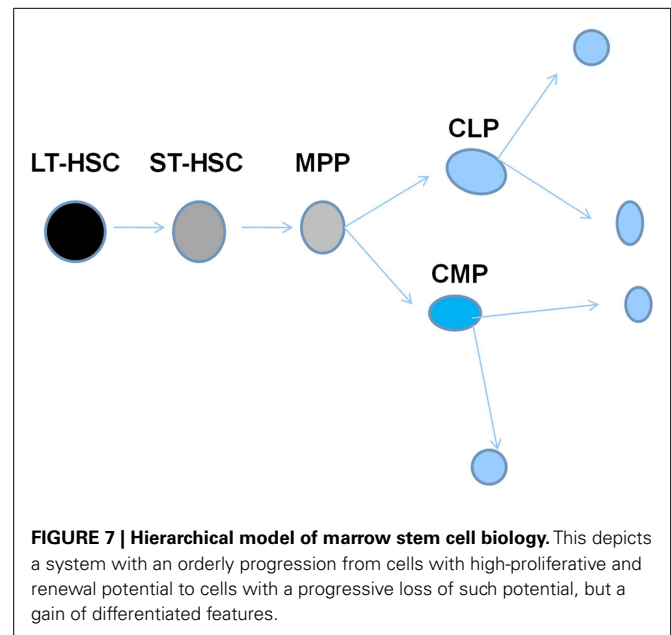
Studies continued to focus on the “holy grail” of stem cell biology; the characterization and isolation of the long-term multilineage repopulating stem cell which had to be defined on a clonal basis. The continued evaluation of transplant potential of cells separated by expression of various surface markers led to a beautiful and rationale model of hematopoiesis. At one point, I think everyone, including myself, was an unrepentant purificator.

Murine marrow cells were labeled with antibodies to different cell surface proteins, separated by FACS and then assessed for long and short-term engraftment and for the lineage choice after engraftment. Long-term hematopoietic stem cells (LT-HSCs) were separated on the basis of lineage negative status and expression of the surface epitopes c-kit and Sca-1 with either Thy 1.1 expression or absence of FLK2. Cells with a multilineage repopulation potential not exceeding 6–8 weeks were then characterized by gain of FLK2 expression. Loss of Thy 1.1 expression with full expression of FLK2 characterized the next differentiation step to the MPP. Common lymphoid and myeloid stem/progenitor cells were then defined by selective expression of IL-7, Fcγ receptor 11/111, and CD34. This created an elegant model of hematopoiesis as outlined in **Figure 7**.

CD150 was subsequently added as a further definer of LT-HSC. With these separations, a very small number of LT-HSC could repopulate an irradiated host mouse. The holy grail appeared to be within reach and this is pretty much the standard hematopoiesis model today.

AN ALTERNATIVE MODEL OF HEMATOPOIESIS OR THE CONTINUUM HERESY

In early studies on engraftment into non-myeloablated mice, murine marrow cells were treated with the cytokines IL-3, IL-6, IL-11 and steel factor in an attempt to increase engraftment levels. However, after 48 h of culture, there was a marked decrease in engraftment capacity (16, 17). Subsequent studies showed, however, that the loss of engraftment was temporary (18) in six separate experiments engraftment returned to at or above baseline



levels with further culture. This was inconsistent with a hierarchical model and suggested a continuum of changing potential. Subsequently, using either whole unseparated murine marrow or highly purified murine marrow Lin[−], rhodamine low, Hoechst low (LRH) stem cells or Lin[−] Sca-1⁺ cells driven through cell cycle by exposure to cytokines, either IL-3, IL-6, IL-11, and steel factor or thrombopoietin, Flt3, and steel factor, we demonstrated that different phenotypic stem cell characteristics were apparent at different points in cycle or times in culture. These changes varied and were generally reversible. Mapping purified stem cells with propidium iodide as they progressed through a cytokine-stimulated cell cycle transit allowed us to estimate phases of cycle for these experiments. We investigated short- and long-term engraftment (18, 19), progenitor numbers (20), homing to marrow (21) and lung (22), expression of adhesion proteins (23, 24) and cell cycle receptors, stem cell surface markers, and cell cycle and other transcriptional regulators (25–27). All were found to vary with cycle transit but at different points in cycle. Further studies showed that differentiation into megakaryocytes and granulocytes occurred at specific cycle times, so-called “hotspots” and these were reversible (28). Formation of epithelial lung cells from engrafted marrow also varied with cell cycle (22). Most recently, we have shown that microvesicle entry into marrow stem cells varies with cell cycle status (29). This is probably another determinant of stem cell fate. Essentially, every biologic parameter which we investigated varied as stem cells progressed through cell cycle under cytokine stimulation. Data on phenotype variation after engraftment into irradiated mice are presented in **Table 1**.

Cytokine cocktails were either: IL-3, IL-6, IL-11, and steel factor or FLT3L, steel factor, and thrombopoietin. Cycle mapping with cytokine-stimulated (IL-3, IL-6, IL-11, and steel factor) purified LRH stem cells showed an initial cycle length of 36–40 h with subsequent cycle occurring every 12 h. G1 phase was estimated at about 18 h and mid-S-phase at about 28–30 h. Studies of

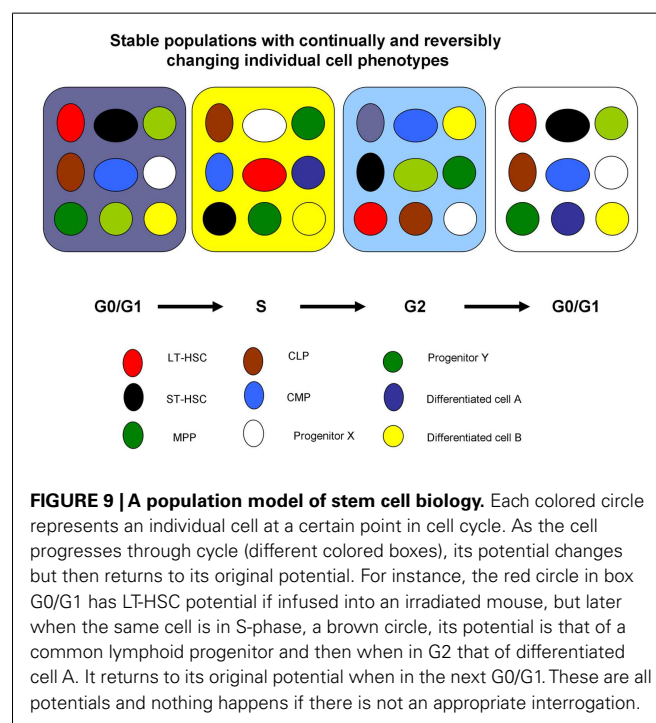
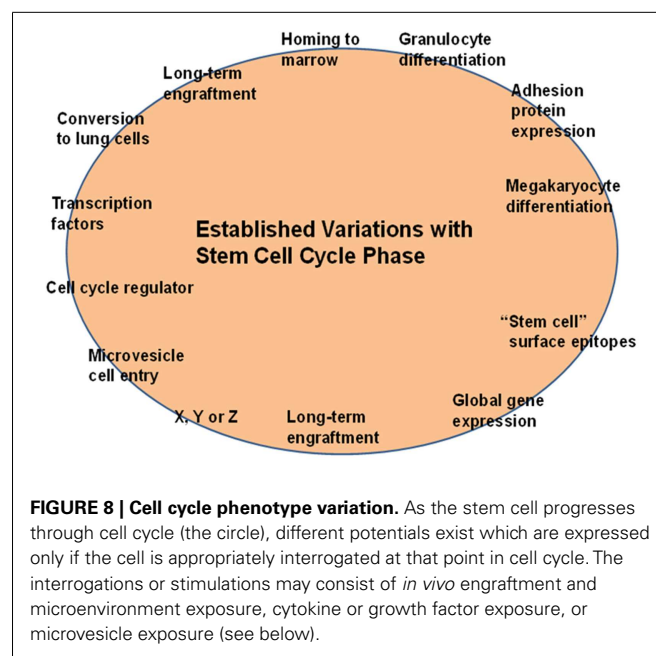
Table 1 | Reversible variations in marrow cell phenotype with cell cycle phase after engraftment into lethally irradiated mice.

Characteristic	Whole marrow	LRH	Lin-Sca-1+
Engraftment	Nadir late S/early G2	Nadir late S/early G2 first and second cycles with recovery in between	
Homing			Nadir late S/G2 second cycle
Increase in lung conversion			Early S-phase
Differentiation into megakaryocyte and proliferative granulocytes		Around G1/S interface	
Differentiation into non-proliferative granulocytes		Late S-phase	

expression of over 40 different genes showed low-level expression of all in LRH cells at isolation and a relatively chaotic variation of expression with cycle transit (25). Adhesion proteins vary but generally drop with cycle transit; CD44 and alpha L increased at 48 h. Studying Lin-Sca-1+ marrow cells during cycle transit under IL-3, IL-6, IL-11, and steel factor stimulation, expression of CD34, CD45R, c-kit, Gata-1, Gata-2 Ikaros, and Fog were stable while Sca-1, Mac-1, c-fms, c-mpl, Tal-1, endoglin, and CD4 showed variation in expression. All showed reversibility except Tal-1, endoglin, and c-mpl. We have also studied LRH cells stimulated by different combinations of cytokines and cloned on a single cell basis (30). Mean cloning efficiency was 31.7% with a range of 8.3–65%. Gross colony morphology and size showed total heterogeneity. Over 100,000 cells per clone were seen at the highest cytokine level. Virtually total heterogeneity as to differentiation phenotype at different points in cycle (0, 18, 32, 40, and 48 h culture) was also demonstrated. There were, however, different patterns of differentiation at different points in cycle; again total individual cellular heterogeneity with population profiles.

These observations, that there were cycle-related reversible changes in stem cell phenotype, suggested a continually changing population of cells consistent with a continuum of cellular potential related to cell cycle phase. This further suggested that while there might be a stable stem cell population, the individual cells or entities in the population were continually changing – shades of Till et al. (2) (see above). A simplified continuum model is presented in **Figure 8** and a more complex model in **Figure 9**.

This model is essentially a model of cell cycle-related continuous change of potentials. It implies that there always will be cohorts of the marrow stem/progenitor population available to respond to any relevant need and the different populations of cells will be continually entering into a responsive window. This is not consistent with the standard hierarchical model of stem cell biology.



CELL CYCLE STATUS OF LONG-TERM ENGRAFTING MULTILINEAGE MARROW STEM CELL – “PAY NO ATTENTION TO THAT MAN BEHIND THE CURTAIN”

A critical consideration as to the biologic relevance of our observations on phenotype change in stem cells as they progress through a cytokine-stimulated cell cycle transit was whether *in vivo* the marrow stem cell is a cycling cell. Much current dogma has it that it is a dormant non-cycling cell, but as we will show, the foundation for this conclusion may have been based on studying the wrong

cell population; the purified LT-HSC. A great deal of mechanistic research has now occurred focusing on the purified LT-HSC. Studies have implicated a myriad of entities as key regulators of hematopoiesis. There have also been a large number of studies attempting to define the hematopoietic stem cell niche employing purified stem cells as a critical tool. However, certain disquieting observations were ignored or not put in the proper context. As noted above, the Ogawa cycle studies (31, 32) indicated that there could not be a straightforward hierarchy, then there were the observations on stem cell plasticity. Much time and effort was wasted here on disputing transdifferentiation versus dedifferentiation, when in fact these studies simply demonstrated that hematopoietic marrow stem cells could be induced to differentiate into non-hematopoietic-cell classes. This of course did not fit with the hematopoietic hierarchy at all and engendered some vigorous attacks from those espousing the conventional hierarchical dogma. Demands that the studies show robustness, be clonal, and not be due to fusion were essentially *ignoratio elenchi* or red herrings (33). Another disquieting fact was ignored. During these separations, the great bulk of marrow long-term repopulating stem cells are lost and the losses are not random. In a non-ablated transplant model, the loss of engraftable stem cells with a LRH purification ranged from 93.6 to 99.2% of what was present in the starting marrow population (34). We have now confirmed these losses in a lethally irradiated mouse model. A critical consideration is that while the final product of purification gives highly purified cells with a specific functional characteristic, the bulk of the stem cells are in the discard fractions. In these fractions, while the percent of stem cells is low, the total number of stem cells is vastly superior to the number seen in the purified fractions.

STEM CELL PURIFICATION AND THE HOLY GRAIL OF SINGLE CELL CLONALITY – A RESEARCH FIELD MISLED

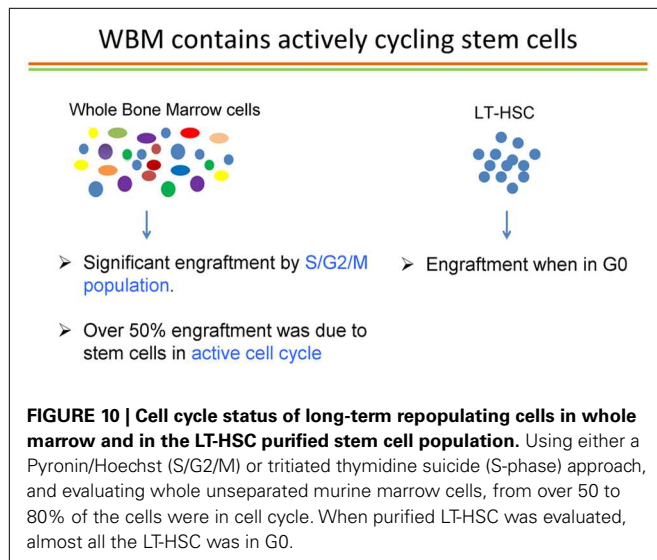
If one enters the descriptor “murine hematopoietic marrow stem cells” into PubMed, one gets over 17,000 hits. Many are not applicable to adult murine marrow stem cells as classically defined; rather they refer to human studies, mesenchymal stem cells, aspects of stem cell plasticity, or other unrelated topics. However, screening these “hits,” there were a large number which referred to aspects of murine adult stem cell biology. These studies involved different purified populations of stem cells as outlined above. In general, the initially published surface phenotype of a functionally defined cell was assumed to hold and functional studies were only rarely carried out. It was assumed that the surface epitope phenotype represented a specific class of stem cells with specific functional characteristics such as long-term or short-term multilineage engraftment or engraftment with differentiation into lymphoid cells. Thus, the vast majority of reported studies of stem cell characteristics: gene expression, cytokine responsiveness, transcriptional regulation, homing, niches, engraftment, and cell cycle status were carried out employing these surrogate phenotypes and assuming stability of these phenotypes. Our continuum studies challenge these concepts. In a similar vein, we have recently reported that the short-term hematopoietic stem cell (ST-HSC, as defined by Lin[−]/Sca-1+/c-kit+/Flk2[−]) was not short-term in our functional experiments involving studies of stem cell homing (35). In addition, as noted above, in studies on highly purified LRH stem

cells, isolated at different points in cell cycle, and grown as single cells in a permissive cytokine cocktail, total heterogeneity of differentiation phenotype was demonstrated (30). Thus, the phenotype was varied and no stability could be inferred. This harks back to the isotope model of Till et al. (2).

We were not the only ones to publish data challenging the conventional concepts of a hierarchical system. Sieburg and colleagues (36) studied 97 individual HSCs in long-term transplantation assays. HSC clones were obtained from unseparated bone marrow (BM) through limiting dilution approaches. Following transplantation into individual hosts, donor-type cells in blood were measured bimonthly and the resulting repopulation kinetics were grouped according to overall shape. Only 16 types of repopulation kinetics were found among the HSC clones even though combinatorially 54 groups were possible. These data were also inconsistent with a straightforward hierarchy. As pointed out to authors in a published correspondence (37), one needs to alter only a few parameters to arrive at the existence of a huge number of stem cell phenotypes which would be most consistent with a continuum model of stem cell biology. Even single cell repopulation assays with highly purified stem cells have been inconsistent with defined stem cells giving rise to an ordered hierarchical system of hematopoiesis. The capacity to isolate a specific stem cell phenotype is of course dependent upon the stability of that phenotype. This in turn is dependent upon the quiescence of the cell under consideration. A cycling cell continually changes phenotype, as we have repeatedly demonstrated, and thus cannot be characterized by a single set of cell surface characteristics. Stability was addressed in more general terms by Montaigne in “Of Repentance” where he states “all things in it are in constant motion; the earth, the rocks of the Caucasus, the pyramids of Egypt, both with the common motion and with their own. Stability itself is nothing but a more languid motion.” These considerations led to a detailed evaluation of the cell cycle status of the murine engrafting marrow stem cell.

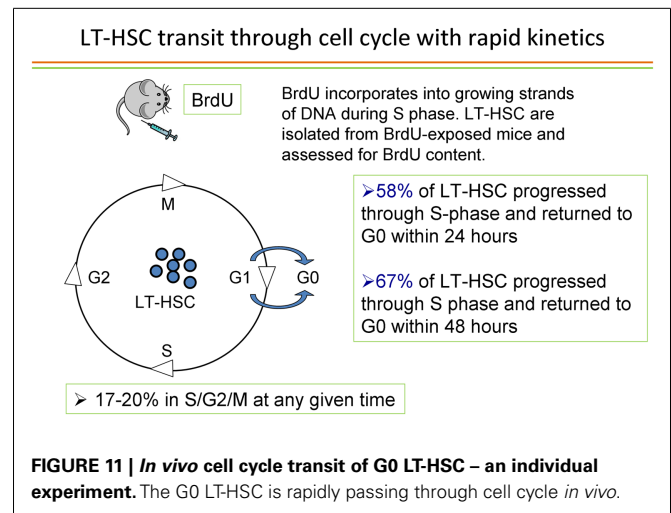
THE CELL CYCLE STATUS OF STEM CELLS – THEY ARE CYCLING!

Passegue and colleagues (38) published elegant studies showing that the long-term repopulating stem cell characterized as Lin[−], c-kit⁺ Sca-1⁺ Flk2[−] only engrafted as a G0 cell. If this was the status of the true marrow stem cells, then our studies of cycle transitioning stem cells could represent an *in vitro* artifact of the culture systems employed. Accordingly, we embarked on a detailed evaluation of the cell cycle status of LT-HSC. With a few outliers, we essentially confirmed the prior studies by Passegue et al. (38) on purified marrow stem cells. We purified LT-HSC into G0, G1, and S/G2/M phases using the supravital dyes Pyronin and Hoechst and then competitively engrafted them into lethally irradiated mice. With a few rare events, essentially all engraftment was found to reside in the G0 compartment of LT-HSC. However, in the review of the literature noted above, we found that essentially all cell cycle studies of engraftable stem cells had been carried out on highly purified stem cells, not on the unseparated whole marrow population. We sought to remedy this oversight by studying unseparated murine marrow cells and determining the cell cycle status of long-term engrafting cells in these cell populations. Accordingly, we separated murine marrow cells into G0,



G1, or S/G2/M populations and then determined long-term multilineage engraftment and secondary engraftment. Over 50% of engraftment was found in the S/G2/M populations of marrow cells. This represented an instantaneous view of cycle status of stem cells and implied that virtually all stem cells must be in cycle (39). We sought to confirm these observations with an alternative method of cycle determination; tritiated thymidine suicide. In this approach, high specific activity tritiated thymidine, a beta emitter, is incubated with marrow cells and, if the cells are synthesizing DNA, the thymidine will be incorporated into the cellular DNA and the cell will then die a radioactive death. The beta particles only exert their activity within the cell and there is no innocent bystander effect. The control marrow cells are incubated with a comparable amount of cold thymidine. At the end of 30 min, a large excess of unlabeled thymidine is added, which inhibits further uptake of the radiolabeled thymidine and the washed cells are then evaluated for long-term multilineage engraftment in a competitive transplant model in lethally irradiated mice. The decrease in engraftment of the tritiated thymidine-treated cells compared to the unlabeled thymidine-treated cells then represents the cell cycle status of these cells. Applying this approach to unseparated B6.SJL marrow cells and then competitively transplanting them into lethally irradiated C57BL/6J mice, we demonstrated that over 70% of the cells had passed through S-phase, thus confirming the active cell cycle status of long-term engrafting cells in normal unseparated murine marrow cells (39). This is summarized in **Figure 10**.

These observations suggested that almost all long-term repopulating stem cells in murine marrow are in cycle. What about the G0 status of engrafting purified LT-HSC? What is the history of these cells? Are they passing through cycle or do they represent a rare permanently quiescent population of cells? In order to address these questions, we utilized *in vivo* BrdU labeling of marrow cells over time. BrdU was administered to B6.SJL mice intraperitoneally (1 mg every 8 h) over 48 h along with BrdU in the drinking water and at different time points, G0 LT-HSC were interrogated for BrdU labeling. BrdU is incorporated into cells



synthesizing DNA and thus provides a cycle passage history for the G0 LT-HSC. At 24 h, 58% of G0 LT-HSC was labeled and at 48 h over 65% were labeled (39). The method and results of a representative experiment is shown in **Figure 11**.

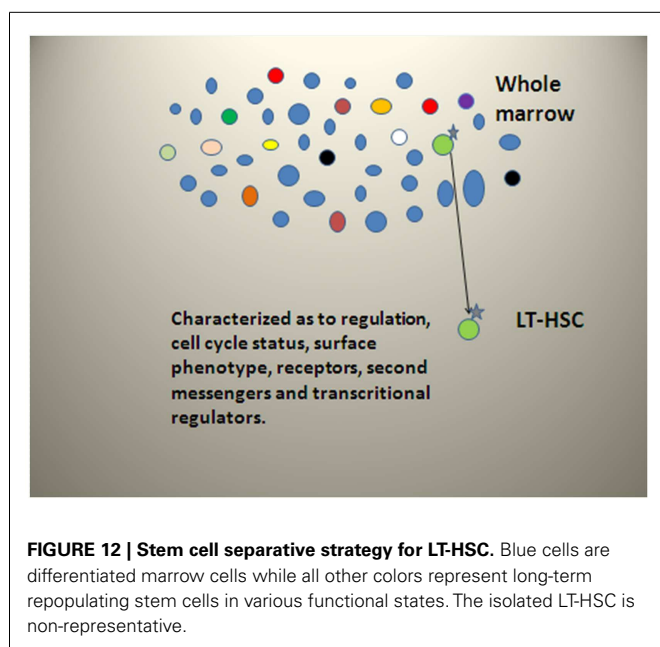
Additional studies ruled out BrdU activation of stem cells into cell cycle. Thus the engraftable LT-HSC is continuously and rapidly transiting cell cycle. This has profound implications to the interpretation of stem cell studies since it indicates that the stem cell phenotype must be continually changing and thus purification of the stem cells is not feasible, rather definition of the stem cell population is the critical issue.

WHY THE DIFFERENCE IN CYCLE STATUS BETWEEN PURIFIED STEM CELLS AND STEM CELLS IN WHOLE UNSEPARATED MARROW? WE FORGOT ABOUT THE DISCARD!

As noted above, we have published studies on purification of LRH stem cells. We showed then that with purification, from 94 to 99% of stem cell capacity was lost (34). We are just now understanding the true significance of these findings. In the course of stem cell purification, almost all of the long-term engraftable marrow stem cells are discarded. Thus, while the purified cells are certainly enriched in stem-like cells, the discarded populations have almost all the stem cells and these cells are cycling. Our data clearly indicate that the vast majority of long-term repopulating stem cells are lost with the separation, which selects out a non-representative dormant cell with long-term repopulating capacity. Current ongoing experiments indicate that most of the proliferating stem cells are in the lineage positive population. This separative strategy is summarized in **Figure 12**.

THEY ALL GOT RHYTHM

All biologic systems have circadian rhythms. They represent basic features of life, but are generally ignored in stem cell studies, because they make stem cell studies very complicated. However, they must be addressed if we are to understand stem cell biology. We have previously carried out relatively limited studies of circadian rhythms of engrafting marrow stem cells and progenitors (40). In these studies, male B6.SJL mice were entrained for 2 weeks

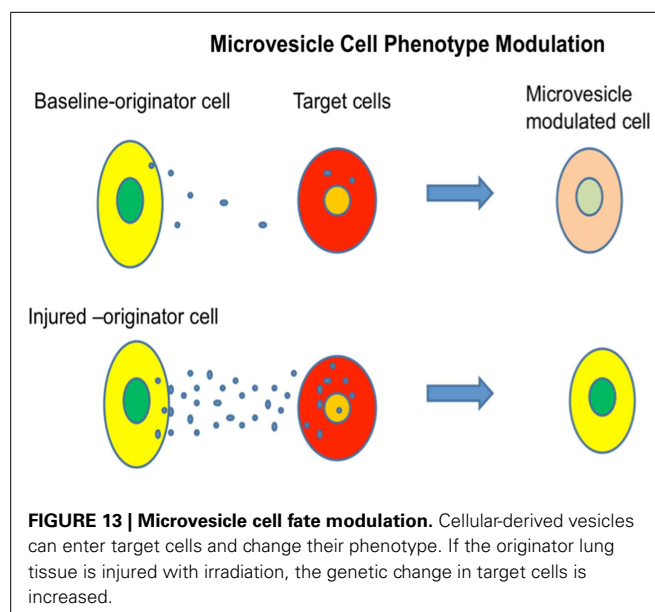


in light dark boxes and then marrow harvested at different circadian times [hours after light onset (HALOs)]. We harvested marrow cells at HALOs 4, 8, 12, 16, 20, and 24. C57BL/6J male hosts at HALO 9 were then subjected to 100 cGy whole body irradiation and injected with 40,000,000 marrow cells from each HALO. Engraftment was then assessed in spleen, marrow, and thymus 10 weeks after cell infusion. In studies carried out in July, there were significant nadirs seen at HALO 8 and HALO 24 with up to fivefold differences between comparative peaks. In separate experiments, we determined that host engraftability showed no circadian rhythms for engraftment. There were progenitor nadirs, HPP-CFC, and total progenitors, at 12 and 24 h. Cycle status of HPP-CFC was determined using tritiated thymidine suicide; increased numbers of HPP-CFC in S-phase were seen at 8, 12, and 24 h.

These data introduce another, usually neglected, variable which needs to be addressed in stem cells studies. We are returning to these methods with regard to stem cell cycle studies.

STEM CELL PLASTICITY (IGNORATIO ELENCHI) AND MICROVESICLES

We have outlined extensive plasticity within the hematopoietic stem/progenitor system above. This was variously interpreted and arguments about transdifferentiation, cell fusion, erroneous cell marking, and quantitative and functional significance ensued. A list of criteria for true plasticity was put forward, which included the necessity of the phenomena being “robust,” clonal, functional, and not due to cell fusion. This was commented on in a perspective in science termed “Ignoratio Elenchi or irrelevant conclusions” (33). This controversy served to halt progress in this area of research and the odor of it still lingers. However, there have now been an overwhelming number of studies indicating that after transplantation of marrow cells, many non-hematopoietic cell types evidenced expression of markers of the donor marrow

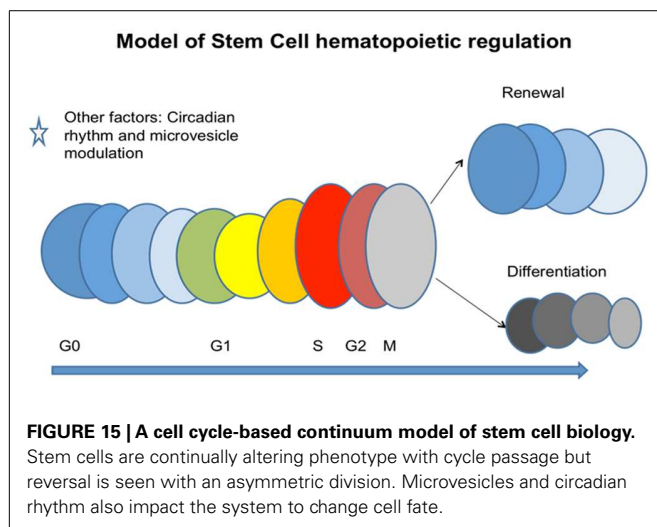
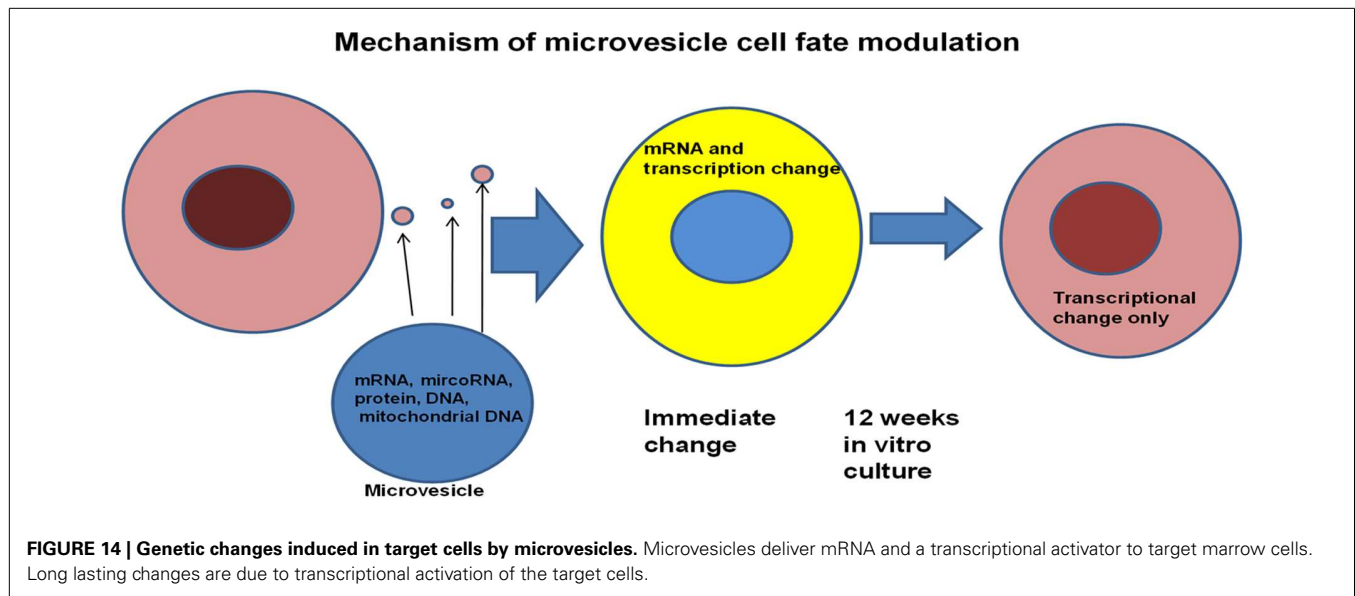


cells. These data were inconsistent with the traditional hierarchical models of hematopoiesis, but fit with the continuum model.

We demonstrated marrow-derived markers in skeletal muscle (41, 42), skin (43), and lung (44) and focused our studies on lung. Utilizing transgenic green-fluorescent protein (GFP) expressing mice as marrow donors, we demonstrated relatively high levels of GFP positive cells in lung, which were further enhanced if host mice were treated with granulocyte colony-stimulating factor after transplantation. Irradiation of host mice was necessary to demonstrate these phenomena.

We investigated mechanisms underlying marrow “transformation” to lung type cells, culturing normal or irradiated lung opposite murine marrow cells, but separated from them by a 0.4 μ m cell impermeable membrane, and then determining whether the marrow cells expressed lung-specific mRNA (45, 46). After 2 or 7 days of co-culture, marrow cells expressed high levels of surfactants A, B, C, and D, Clara cell-specific protein, or aquaporin 5 and this expression was significantly higher in marrow cultured across from irradiated lung as compared to non-irradiated lung. Lung-conditioned media could elicit the same genetic changes in marrow cultured in the conditioned media and it was then determined that the active principle could be spun down by ultracentrifugation, the pellet containing large numbers of microvesicles (Figure 13).

Further studies showed that the genetic change was dependent upon microvesicles entering target marrow cells. All classes of marrow cells imbibed microvesicles and exposure to lung microvesicles were shown to increase the capacity of modulated marrow cells, engrafted into irradiated mice, to “convert” to epithelial lung cells approximately twofold. Microvesicles themselves contained mRNA, protein, microRNA, and mitochondrial and genomic DNA. The microvesicles also expressed many surface proteins including adhesion proteins. Mechanistic studies employing rat/mouse hybrid co-cultures with rat- and mouse-specific primers for mRNA for surfactants B and C showed immediate transfer of originator cell mRNA and also induction of target



cell mRNA (46). However, the originator cell mRNA disappeared rapidly with time in cytokine-supported liquid culture, while the target cell mRNA persisted for out to 12 weeks in culture. Transplanted cells also showed expression of lung-specific mRNA out to 6 weeks (as far as tested) in marrow, thymus, liver, and lung. These studies indicated that a persistent epigenetic transcriptional change had occurred (Figure 14).

Very similar studies were carried out with rat liver and mouse marrow co-culture evaluating rat- and mouse-specific mRNA for albumin, with the same results. It was also shown that murine brain and murine heart induced tissue-specific mRNA in target marrow cells. Other work has indicated that mesenchymal-derived microvesicles could mediate *in vivo* repair of renal damage (47). Altogether these data suggest that microvesicles may represent a general biologic cell phenotype modulating mechanism, adding further complexity to marrow stem cell models. A general model encompassing many of the above noted variables is presented in Figure 15.

TO THE LAST PLACE OF DECIMALS

There was a time in the late 1800s when physicists were “assured of certain certainties” and felt that essentially all the basic aspects of physics had been elucidated and that “our future discoveries must be looked for in the sixth place of decimals” (48). Then came quantum mechanics which changed everything. In a similar fashion, many in the hematopoietic stem cell field appear to feel that the stem cell is close to final definition with progressive progress in purification and that next steps simply are going to the sixth place of decimals. Rather, we think that we are entering the area of quantum stemonics where a true understanding of stem cell biology beckons.

ACKNOWLEDGMENTS

This work was supported by the National Center for Research Resources (NCRR) and the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) through grant number 8P20GM103468-04 and by the National Heart, Lung and Blood Institutes (NHLBI) of NIH through grant number R01HL103726. We thank Laura Bangs, Elaine Papa, and Paula Salisbury for assistance in the preparation of this manuscript.

REFERENCES

1. Till JE, McCulloch EA. A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. *Radiat Res* (1961) **14**:213–22. doi:10.2307/3570892
2. Till JE, McCulloch EA, Siminovitch L. A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. *Proc Natl Acad Sci U S A* (1964) **51**:29–36. doi:10.1073/pnas.51.1.29
3. Magli MC, Iscove NN, Odartchenko N. Transient nature of early hematopoietic spleen colonies. *Nature* (1982) **295**(5849):527–9. doi:10.1038/295527a0
4. Bradley TR, Metcalf D. The growth of mouse bone marrow cells in vitro. *Aust J Exp Biol Med Sci* (1966) **44**(3):287–99. doi:10.1038/icb.1966.28
5. Pluznik DH, Sachs L. The cloning of normal “mast” cells in tissue culture. *J Cell Physiol* (1965) **66**(3):319–24. doi:10.1002/jcp.1030660309
6. Iscove NN, Sieber F. Erythroid progenitors in mouse bone marrow detected by macroscopic colony formation in culture. *Exp Hematol* (1975) **3**(1):32–43.

7. Gregory CJ. Erythropoietin sensitivity as a differentiation marker in the hematopoietic system: studies of three erythropoietic colony responses in culture. *J Cell Physiol* (1976) **89**(2):289–301. doi:10.1002/jcp.1040890212
8. Long MW, Gragowski LL, Heffner CH, Boxer LA. Phorbol diesters stimulate the development of an early murine progenitor cell. The burst-forming unit-megakaryocyte. *J Clin Invest* (1985) **76**(2):431–8. doi:10.1172/JCI111990
9. Suda T, Suda J, Ogawa M. Single-cell origin of mouse hematopoietic colonies expressing multiple lineages in variable combinations. *Proc Natl Acad Sci U S A* (1983) **80**(21):6689–93. doi:10.1073/pnas.80.21.6689
10. Bradley TR, Hodgson GS. Detection of primitive macrophage progenitor cells in mouse bone marrow. *Blood* (1979) **54**(6):1446–50.
11. Spangrude GJ, Heimfeld S, Weissman IL. Purification and characterization of mouse hematopoietic stem cells. *Science* (1988) **241**(4861):58–62. doi:10.1126/science.2898810
12. Morrison SJ, Weissman IL. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity* (1994) **1**(8):661–73. doi:10.1016/1074-7613(94)90037-X
13. Christensen JL, Weissman IL. Flk-2 is a marker in hematopoietic stem cell differentiation: a simple method to isolate long-term stem cells. *Proc Natl Acad Sci U S A* (2001) **98**(25):14541–6. doi:10.1073/pnas.261562798
14. Kiel MJ, Yilmaz OH, Iwashita T, Yilmaz OH, Terhorst C, Morrison SJ. SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* (2005) **121**(7):1109–21. doi:10.1016/j.cell.2005.05.026
15. Osawa M, Hanada K, Hamada H, Nakauchi H. Long-term lymphohematopoietic reconstitution by a single CD34-low/negative hematopoietic stem cell. *Science* (1996) **273**(5272):242–5. doi:10.1126/science.273.5272.242
16. Peters SO, Kittler EL, Ramshaw HS, Quesenberry PJ. Ex vivo expansion of murine marrow cells with interleukin-3 (IL-3), IL-6, IL-11, and stem cell factor leads to impaired engraftment in irradiated hosts. *Blood* (1996) **87**(1):30–7.
17. Peters SO, Kittler EL, Ramshaw HS, Quesenberry PJ. Murine marrow cells expanded in culture with IL-3, IL-6, IL-11, and SCF acquire an engraftment defect in normal hosts. *Exp Hematol* (1995) **23**(5):461–9.
18. Habibian HK, Peters SO, Hsieh CC, Wu J, Vergilis K, Grimaldi CI, et al. The fluctuating phenotype of the lymphohematopoietic stem cell with cell cycle transit. *J Exp Med* (1998) **188**(2):393–8. doi:10.1084/jem.188.2.393
19. Peters SO, Habibian HK, Vergilis K, Quesenberry PJ. Effects of cytokines on stem cell engraftment depends on time of evaluation post-marrow-infusion. *Int J Hematol* (1999) **70**(2):112–8.
20. Colvin GA, Lambert JF, Moore BE, Carlson JE, Dooner MS, Abedi M, et al. Intrinsic hematopoietic stem cell/progenitor plasticity: inversions. *J Cell Physiol* (2004) **199**(1):20–31. doi:10.1002/jcp.10436
21. Cerny J, Dooner M, McAuliffe C, Habibian H, Stencil K, Berrios V, et al. Homing of purified murine lymphohematopoietic stem cells: a cytokine-induced defect. *J Hematother Stem Cell Res* (2002) **11**(6):913–22. doi:10.1089/152581602321080574
22. Dooner MS, Aliotta JM, Pimentel J, Dooner GJ, Abedi M, Colvin G, et al. Conversion potential of marrow cells into lung cells fluctuates with cytokine-induced cell cycle. *Stem Cells Dev* (2008) **17**(2):207–19. doi:10.1089/scd.2007.0195
23. Becker PS, Nilsson SK, Li Z, Berrios VM, Dooner MS, Cooper CL, et al. Adhesion receptor expression by hematopoietic cell lines and murine progenitors: modulation by cytokines and cell cycle status. *Exp Hematol* (1999) **27**(3):533–41. doi:10.1016/S0301-472X(98)00037-X
24. Berrios VM, Dooner GJ, Nowakowski G, Frimberger A, Valinski H, Quesenberry PJ, et al. The molecular basis for the cytokine-induced defect in homing and engraftment of hematopoietic stem cells. *Exp Hematol* (2001) **29**(11):1326–35. doi:10.1016/S0301-472X(01)00734-2
25. Dooner GJ, Colvin GA, Dooner MS, Johnson KW, Quesenberry PJ. Gene expression fluctuations in murine hematopoietic stem cells with cell cycle progression. *J Cell Physiol* (2008) **214**(3):786–95. doi:10.1002/jcp.21273
26. Reddy GP, McAuliffe CI, Pang L, Quesenberry PJ, Bertoncello I. Cytokine receptor repertoire and cytokine responsiveness of Ho(dull)/Rh(dull) stem cells with differing potentials for G1/S phase progression. *Exp Hematol* (2002) **30**(7):792–800. doi:10.1016/S0301-472X(02)00814-7
27. Quesenberry PJ, Dooner GJ, Totto MD, Colvin GA, Johnson K, Dooner MS. Expression of cell cycle-related genes with cytokine-induced cell cycle progression of primitive hematopoietic stem cells. *Stem Cells Dev* (2010) **19**(4):453–60. doi:10.1089/scd.2009.0283
28. Colvin GA, Dooner MS, Dooner GJ, Sanchez-Guijo FM, Demers DA, Abedi M, et al. Stem cell continuum: directed differentiation hotspots. *Exp Hematol* (2007) **35**(1):96–107. doi:10.1016/j.exphem.2006.09.005
29. Aliotta JM, Lee D, Puente N, Faradyan S, Sears EH, Amaral A, et al. Progenitor/stem cell fate determination: interactive dynamics of cell cycle and microvesicles. *Stem Cells Dev* (2012) **21**(10):1627–38. doi:10.1089/scd.2011.0550
30. Colvin GA, Berz D, Liu L, Dooner MS, Dooner G, Pascual S, et al. Heterogeneity of non-cycling and cycling synchronized murine hematopoietic stem/progenitor cells. *J Cell Physiol* (2012) **222**(1):57–65. doi:10.1002/jcp.21918
31. Suda J, Suda T, Ogawa M. Analysis of differentiation of mouse hematopoietic stem cells in culture by sequential replating of paired progenitors. *Blood* (1984) **64**(2):393–9.
32. Suda T, Suda J, Ogawa M. Disparate differentiation in mouse hematopoietic colonies derived from paired progenitors. *Proc Natl Acad Sci U S A* (1984) **81**(8):2520–4. doi:10.1073/pnas.81.8.2520
33. Quesenberry PJ, Dooner G, Dooner M, Abedi M. Developmental biology: ignoratio elenchi: red herrings in stem cell research. *Science* (2005) **308**(5725):1121–2. doi:10.1126/science.1104432
34. Nilsson SK, Dooner MS, Tiarks CY, Weier HU, Quesenberry PJ. Potential and distribution of transplanted hematopoietic stem cells in a nonablated mouse model. *Blood* (1997) **89**(11):4013–20.
35. Liu L, Papa EF, Dooner MS, Machan JT, Johnson KW, Goldberg LR, et al. Homing and long-term engraftment of long- and short-term renewal hematopoietic stem cells. *PLoS One* (2012) **7**(2):e31300. doi:10.1371/journal.pone.0031300
36. Sieburg HB, Cho RH, Dykstra B, Uchida N, Eaves CJ, Muller-Sieburg CE. The hematopoietic stem compartment consists of a limited number of discrete stem cell subsets. *Blood* (2006) **107**(6):2311–6. doi:10.1182/blood-2005-07-2970
37. Roeder I, Loeffler M, Quesenberry PJ, Colvin GA, Lambert JF. Quantitative tissue stem cell modeling. *Blood* (2003) **102**(3):1143–4. doi:10.1182/blood-2003-04-1158
38. Passegue E, Wagers AJ, Giuriato S, Anderson WC, Weissman IL. Global analysis of proliferation and cell cycle gene expression in the regulation of hematopoietic stem and progenitor cell fates. *J Exp Med* (2005) **202**(11):1599–611. doi:10.1084/jem.20050967
39. Goldberg LR, Dooner MS, Johnson KW, Papa EF, Pereira MG, Del Tatto M, et al. The murine long-term multi-lineage renewal marrow stem cell is a cycling cell. *Leukemia* (2013) doi:10.1038/leu.2013.252
40. D'Hondt L, McAuliffe C, Damon J, Reilly J, Carlson J, Dooner M, et al. Circadian variations of bone marrow engraftability. *J Cell Physiol* (2004) **200**(1):63–70. doi:10.1002/jcp.20032
41. Abedi M, Greer DA, Colvin GA, Demers DA, Dooner MS, Harpel JA, et al. Robust conversion of marrow cells to skeletal muscle with formation of marrow-derived muscle cell colonies: a multifactorial process. *Exp Hematol* (2004) **32**(5):426–34. doi:10.1016/j.exphem.2004.02.007
42. Abedi M, Greer DA, Foster BM, Colvin GA, Harpel JA, Demers DA, et al. Critical variables in the conversion of marrow cells to skeletal muscle. *Blood* (2005) **106**(4):1488–94. doi:10.1182/blood-2005-01-0264
43. Badiavas EV, Abedi M, Butmarc J, Falanga V, Quesenberry P. Participation of bone marrow derived cells in cutaneous wound healing. *J Cell Physiol* (2003) **196**(2):245–50. doi:10.1002/jcp.10260
44. Aliotta JM, Keaney P, Passero M, Dooner MS, Pimentel J, Greer D, et al. Bone marrow production of lung cells: the impact of G-CSF, cardiotoxin, graded doses of irradiation, and subpopulation phenotype. *Exp Hematol* (2006) **34**(2):230–41. doi:10.1016/j.exphem.2005.11.007
45. Aliotta JM, Sanchez-Guijo FM, Dooner GJ, Johnson KW, Dooner MS, Greer KA, et al. Alteration of marrow cell gene expression, protein production, and engraftment into lung by lung-derived microvesicles: a novel mechanism for phenotype modulation. *Stem Cells* (2007) **25**(9):2245–56. doi:10.1634/stemcells.2007-0128
46. Aliotta JM, Pereira M, Johnson KW, de Paz N, Dooner MS, Puente N, et al. Microvesicle entry into marrow cells mediates tissue-specific changes in mRNA by direct delivery of mRNA and induction of transcription. *Exp Hematol* (2010) **38**(3):233–45. doi:10.1016/j.exphem.2010.01.002
47. Gatti S, Bruno S, Deregibus MC, Sordi A, Cantaluppi V, Tetta C, et al. Microvesicles derived from human adult mesenchymal stem cells protect against ischemia-reperfusion-induced acute and chronic kidney injury. *Nephrol Dial Transplant* (2011) **26**(5):1474–83. doi:10.1093/ndt/gfr015

48. Michelson A. *Light Waves and Their Uses*. Chicago: The University of Chicago Press (1903).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 August 2013; paper pending published: 26 September 2013; accepted: 11 March 2014; published online: 04 April 2014.

Citation: Quesenberry PJ, Goldberg L, Aliotta J and Dooner M (2014) Marrow hematopoietic stem cells revisited: they exist in a continuum and are not defined by

standard purification approaches; then there are the microvesicles. *Front. Oncol.* 4:56. doi: 10.3389/fonc.2014.00056

This article was submitted to *Cancer Genetics*, a section of the journal *Frontiers in Oncology*.

Copyright © 2014 Quesenberry, Goldberg, Aliotta and Dooner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.