bjh research paper

Inferring relative numbers of human leucocyte genome replications

Michelle Chu,[1] Kimberly D. Siegmund,[2] Qian-Lin Hao,[3] Gay M. Crooks,[3] Simon Tavaré[4,5] and Darryl Shibata[1]

[1]Departments of Pathology and [2]Preventive Medicine, University of Southern California Keck School of Medicine, [3]Division of Immunology and Bone Marrow Transplant, Childrens Hospital Los Angeles, [4]Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA, and [5]Department of Oncology, University of Cambridge, Cambridge, UK

Received 27 December 2007; accepted for publication 5 February 2008

Correspondence: Dr D. Shibata, Department of Pathology, University of Southern California Keck School of Medicine, 1200 N. State Street, Unit I Room 2428, Los Angeles, CA 90033, USA.
E-mail: dshibata@usc.edu

Summary

Genome duplication inevitably results in replication errors. *A priori*, the more times a genome is copied, the greater the average number of replication errors. This principle could be used to 'count' mitotic divisions. Although somatic mutations are rare, cytosine methylation is also copied after DNA replication, but measurably increases with aging at certain CpG rich sequences in mitotic tissues, such as the colon. To further test whether such age-related methylation represents replication errors, these CpG rich 'clock' sequences were measured in leucocytes. Leucocytes within an individual have identical chronological ages (time since birth) but their mitotic ages (numbers of divisions since the zygote) may differ. Neutrophils, B-lymphocytes, and red cell progenitors are produced from relatively quiescent stem cells throughout life, but T-lymphocyte production largely ceases after puberty when the thymus disappears. However, T-lymphocyte genomes may continue to replicate throughout life in response to immunological stimulation. Consistent with this biology, clock methylation significantly increased with aging for T-lymphocyte genomes, but no significant increase was measured in other cell populations. Moreover, this methylation was greater in genomes isolated from their corresponding neoplastic populations. These studies tentatively support the hypothesis that methylation at certain CpG rich sequences in leucocytes could record their mitotic ages.

Keywords: genealogy, development, aging, haematology, epigenetics.

Cell division occurs throughout life and the numbers of divisions are likely to have important roles for normal and abnormal aging. However, measurements of human mitotic ages (numbers of divisions since the zygote) are limited because many experimental interventions that mark and follow cell fates are impractical. Because genome duplication inevitably results in replication errors, one approach to estimate numbers of divisions is to analyze the historical information contained within genomes. Telomere length measurements are commonly used to estimate numbers of somatic cell divisions because chromosome ends shorten from an 'end-replication' problem (Brummendorf & Balabanov, 2006). A more general approach analyzes the sequences between telomeres. Genome comparisons are usually employed to reconstruct ancestry between species or populations (Bromham & Penny, 2003; Hedges & Kumar, 2003). On average, the greater the numbers of divisions, the greater the numbers of differences between two genomes.

It may be possible to translate molecular phylogenetic methods to somatic cell aging because somatic replication errors also occur. All genomes within an individual are copies of the first genome in the zygote, related by an ancestral tree that starts from the zygote and ends with present day cells. Although somatic cells have identical chronological ages (time since birth), they may have different mitotic ages and numbers of replication errors because mitotic activity may differ between cells.

Mutations occur too rarely to function effectively as somatic cell 'molecular clocks'. For example, mutation frequencies are about one per million bases in cancer genomes (Sjoblom *et al*, 2006). However, it may be more practical to trace ancestry by substituting instead the 5′ to 3′ order of covalent epigenetic modifications, such as cytosine methylation at CpG dinucleotides (Shibata & Tavaré, 2006). Methylation patterns are usually copied after DNA replication (Holliday, 1987; Brero

First published online 10 April 2008
doi:10.1111/j.1365-2141.2008.07142.x © 2008 The Authors Journal Compilation © 2008 Blackwell Publishing Ltd,

et al, 2006), but potentially with less fidelity because methylation measurably accumulates at certain CpG sites with age in mitotic human tissues (Issa, 2000). Similar to sequences, 5′ to 3′ methylation patterns can be measured by bisulphite sequencing. Unlike sequences, whose germline may differ between individuals, most CpG sites 'start' unmethylated because most methylation is removed early in development before implantation (Morgan *et al*, 2005). Therefore, it may be possible to observe 'serially' how tissue methylation patterns change by examining differently aged individuals. Cells with greater mitotic ages should contain genomes with greater average numbers of methylated sites or replication errors at certain CpG rich regions.

Translating molecular phylogeny to a somatic cell tree requires the collection of empirical evidence because, *a priori*, replication errors may occur rarely. Some methylation is programmed and associated with transcriptional silencing (Jones & Laird, 1999; Bird, 2002). Programmed methylation is probably directed at genes important for development and differentiation, and methylation patterns in these genes should be similar between cells of a given phenotype. However, only some genes are expressed in a cell type, and methylation within non-expressed or poorly-expressed genes may be dominated by random replication errors, which are defined here as unprogrammed changes in *de novo* methylation and demethylation associated with cell division.

Unprogrammed *de novo* methylation and demethylation should occur, by default, in many cell types. To test this hypothesis, methylation at several CpG rich sequences or 'tags' (Fig 1) were examined in multiple human tissues. Age-related increases in tag methylation were observed (Fig 2) in mitotic epithelium from the colon (Yatabe *et al*, 2001), small intestines (Kim *et al*, 2005a), and endometrium (Kim *et al*, 2005b). Consistent with a correlation between numbers of methylated sites and replication errors, average endometrial tag methylation increased with age but was constant after the menopause, when epithelial division largely ceases. Tag changes do not appear to simply reflect chronological age because average methylation was low in fetal or infant brain, and higher but

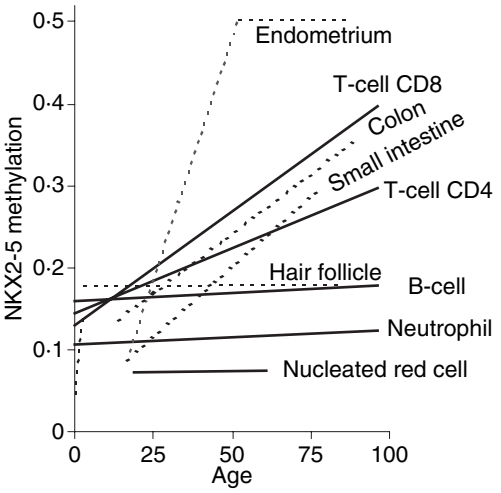

Fig 2. *NKX2-5* tag methylation trends with aging for human leucocytes and previously published studies of colon (Yatabe *et al*, 2001), small intestines (Kim *et al*, 2005a), endometrium (Kim *et al*, 2005b), and hair (Kim *et al*, 2006). Age-related increases in methylation were observed for the epithelial tissues, except for hair, which also appears to exhibit stem cell latency, like HSCs. All human genealogies have common origins (the zygote), and replication errors (methylation) may record relative numbers of divisions in different tissues during aging.

constant in the adult brain, which lacks significant mitotic activity (Chu *et al*, 2007). Of note, cell division appears to be required for *de novo* methylation (Velicescu *et al*, 2002). Hair follicle epithelium is highly mitotic but follicle tag methylation also did not increase after infancy (Kim *et al*, 2006). The lack of a measurable increase in hair follicle mitotic age is consistent with relatively quiescent hair stem cells (see *Discussion*).

Human leucocytes (Shizuru *et al*, 2005) provide another opportunity to test whether unprogrammed tag methylation and demethylation can effectively function as somatic cell molecular clocks. Each type of leucocyte has a characteristic genealogy (Fig 3A) but all originate from pluripotent haematopoietic stem cells (HSCs). Leucocytes representing different

NKX2–5 (8 CpG sites, Chr 5q34)

ggagaTTaggaacttttTtgtTTTa**CG**C**G**C**G**tttgttTttg**CG**Ta**CG**ggagagtttgtgg**CG**g**CG**attatgTag**CG**tgTaatgagtgatTTtgTagTTtggtgtT

BGN (9 CpG sites, Chr Xq28)

TTTaggagtgagtagTtgTttt**CG**gtT**CG**T**CG**gaTaTaT**CG**gaTagataga**CG**tg**CG**ga**CG**gTTTaTTaTTTTagTT**CG**TTaaTtagtTagTTtg
CGTTtgg**CG**TTtTTTTtTtTTaggtagggTtggT

SOX10 (20 CpG sites, Chr 22q13·1)

ggTaaggtTaagaaggagTagTagga**CG**g**CG**agg**CG**ga**CG**atgaTaagttTTT**CG**tgtgTatT**CG**C**G**aggT**CG**tTagTTaggtgTtTag**CG**gTta**CG**aTt
gga**CG**TtggtgTTTatgTT**CG**tg**CG**C**G**tTaa**CG**g**CG**TTagTaaaagTaagT**CG**Ta**CG**tTaag**CG**gTTTatgaa**CG**TTttTatggtgtgggTtTaggTag

BMI1 (7 CpG sites, Chr 10p11·23)

GtTttttaggaTttggtgggaaTttaaagatgTtt**CG**aaaagTTtggaaagagTaaTtTagttgtaaTttggggaagttaa**CG**atT**CG**TT**CG**agtgTagaggaaaa
TTagaaa**CG**TTttgTTttTagTtgaaTTaT**CG**tttgtg**CG**agTtggatgtTTttttTagtagaa

Fig 1. 'Clock' tags used in this study. Sequences are after bisulphite treatment that chemically converts C to T, but methyl-C is not changed. CpG sites are bold and 'Ts' indicate converted 'Cs' at non-CpG sites. PCR primer sites are underlined. *NKX2-5*, *BGN*, and *SOX10* clock tags were previously used to study human tissues (Yatabe *et al*, 2001; Kim *et al*, 2006).

© 2008 The Authors
Journal Compilation © 2008 Blackwell Publishing Ltd,

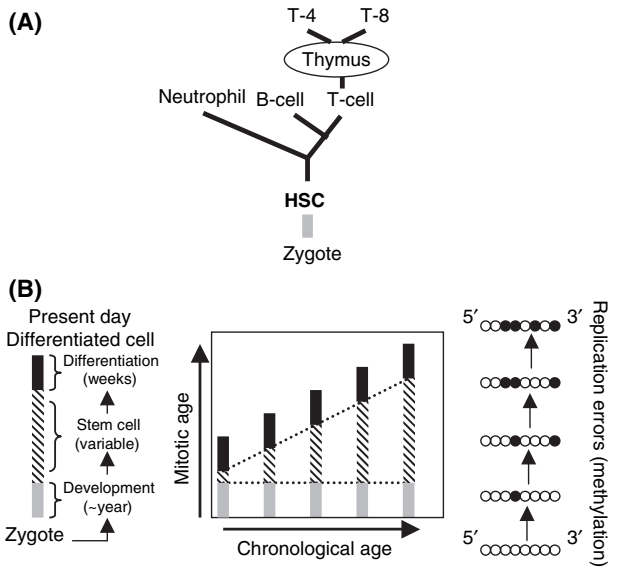

Fig 3. A haematopoietic somatic cell ancestral tree. (A) A haematopoietic tree starts from the zygote, with the HSC a common ancestor of neutrophils, red cells and lymphocytes. The HSC is a relatively rare, pluripotent stem cell capable of restoring haematopoiesis after transplantation. Numbers of bone marrow progenitors increase progressively during differentiation. T-cell maturation occurs outside the bone marrow in the thymus. (B)Somatic cell genealogy can be divided into three phases: development from the zygote to a stem cell, a stem cell phase, and differentiation. Development and differentiation are programmed and therefore restricted to certain times and numbers of divisions. Therefore, only the stem cell phase can have a variable time interval, typically between days to decades. Mitotic age is the total numbers of divisions since the zygote, and the mitotic age of a differentiated cell largely reflects the mitotic activity of its stem cell. An increase in the mitotic age of differentiated cells with chronological age implies that their stem cells divide throughout life. Mitotic age can potentially be measured using a molecular clock approach – the greater the numbers of cell divisions, the greater the numbers of replication errors. Certain CpG rich regions appear to function as epigenetic molecular clocks because are unmethylated early in development and randomly accumulate methylation in mitotic epithelium like the human intestines (open and filled circles represent the 5′ to 3′ order of unmethylated and methylated CpG sites, respectively).

parts of the haematopoietic ancestral tree can be purified and their tags measured. Neutrophils are short-lived, terminally differentiated leucocytes that no longer divide. Red cell progenitors also stop dividing and differentiate after a few weeks into enucleated red cells that circulate in the blood. Lymphocytes originate from HSCs but may have much greater mitotic ages because they are longer-lived and divide in response to immunological stimulation after their release from the bone marrow (Sprent & Tough, 1994). B-cells are produced throughout life (Nunez *et al*, 1996), but fewer T-cells are produced with aging because maturation is dependent on the thymus, which regresses after puberty. T-cell numbers do not decline with age, suggesting they may survive a lifetime and that peripheral divisions are important after thymic regression (Linton & Dorshkind, 2004; Goronzy & Weyand, 2005). Here we test whether tag methylation records the basic genealogy of human leucocytes.

Materials and methods

Specimens

Peripheral blood was obtained from normal volunteers or excess clinical specimens with normal blood counts from the Los Angeles County–University of Southern California Medical Center or the Norris Cancer Center. Bone marrow or cord blood specimens were from excess clinical specimens. Studies were approved by the Institutional Review Board. Specific cell types were purified by flow cytometry. Neutrophil gates were based on forward and side scatter. Other leucocytes were sorted on scatter and labelled antibodies (B-lymphocytes, $CD45^+$,$CD20^+$; T-lymphocytes, $CD3^+$,$CD4^+$, or $CD3^+$,$CD8^+$; haematopoietic progenitors, low side scatter and $CD34^+$; and marrow nucleated red cells, $CD45^-$, $GYPA^+$). Leukaemia blasts were sorted on scatter and labelled antibodies (pre-B ALL, CD45dim, $CD10^+$; T-cell ALL, CD45dim, $CD7^+$; AML, CD45dim, and generally $CD34^+$ and $CD33^+$ or $CD117^+$). Between 3000 and 20 000 cells were collected for each cell type.

Lymphoid cultures and colony forming units (CFU) were grown from cord blood and adult sources (normal bone marrow of a 20-year-old subject, and G-CSF mobilized peripheral blood mononuclear cells from a 27-year-old subject) as previously described (Hao *et al*, 1998).

Methylation analysis

DNA was isolated and bisulphite treated as previously described (Yatabe *et al*, 2001). Briefly, cells were digested in a Tris–HCl buffer with Proteinase-K at 56°C for 2 h, and then boiled for 5 min. Bisulphite treatment (12 h at 50°C) used an agarose bead method, which prevents loss of small amounts of DNA. Approximately 10% of the bisulphite treated bead was amplified for 42 polymerase chain reaction (PCR) cycles in duplicate, with PCR primers specific for the *NKX2-5* (eight CpG sites, also known as *CSX*), *BGN* (nine CpG sites), *SOX10* (20 CpG sites), and *BMI1* (six CpG sites) clock tags (Fig 1). The duplicate PCR products were mixed prior to cloning (TOPO TA Cloning kit; Invitrogen, Carlsbad, CA, USA). Eight bacterial clones per sample were sequenced. Tags with incomplete bisulphite conversion ('C' at non-CpG sites) were discarded from the analysis. The methylation of each tag was the average of all its sites, with average methylation for each specimen the average of eight tags. Statistical comparisons, unless otherwise noted, used a two-sided *t*-test, with significance at $P < 0{\cdot}05$. Comparisons were between individual tags except for clonal specimens (leukaemia and CFUs), where specimen tag averages were compared.

Mitotic age is inferred by counting numbers of replication errors, or methylated CpG sites. The greater the number of divisions, the greater the average number of replication errors

© 2008 The Authors

or methylated CpG sites (Shibata & Tavaré, 2006). We assumed that the tags started out unmethylated early in life, consistent with the observation that most CpG islands are unmethylated at birth (Bird, 2002), and experimental evidence that our tags exhibited low methylation in fetal tissues (Kim *et al*, 2006; Chu *et al*, 2007). Both forward (methylation) and backward (demethylation) errors are possible, but the process is initially dominated by forward errors because the tags start out unmethylated. The stochastic nature of mutation will result in discordance between mitotic ages and numbers of methylated sites. Some cells with greater mitotic ages will have less methylation than cells with lower mitotic ages, and visa versa. Because individual tags are relatively uninformative, eight tags were sampled per specimen.

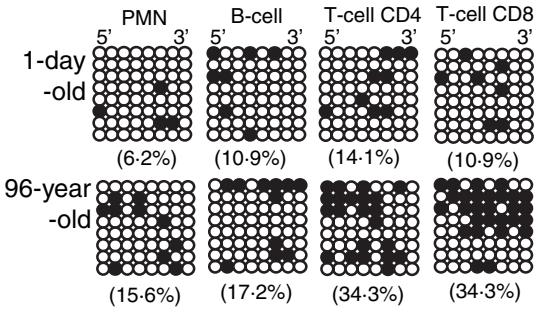

Fig 4. Tags from subjects aged 1 d and 96 years. Each *NKX2-5* tag contains eight CpG sites, arranged horizontally in a 5′ to 3′ order, with eight tags sampled from each cell type. Open circles are unmethylated CpG sites and filled circles are methylated sites. The average methylation of the eight tags is indicated in parentheses. PMN, polymorphonuclear cells.

Results

The methodology assumes that CpG rich tags (Fig 1) start unmethylated, consistent with the genome-wide loss of methylation that occurs before implantation (9). The tags (*NKX2-5*, *BGN*, and *SOX10*) previously exhibited 'clock-like' age-related methylation in human intestines and uteri (Fig 2). The *BGN* locus, on the X-chromosome, was examined only in males where there is only a single haplotype per cell. During subsequent cell division, both unprogrammed methylation and demethylation replication errors were possible, but net change favoured methylation, especially because the tags started unmethylated (Fig 3B).

The measurement of a single tag from a single cell is relatively uninformative because of the stochastic nature of replication errors and low methylation error rates (estimated at about 10^{-5} per CpG site per division for *NKX2-5* and *BGN* in the colon (Yatabe *et al*, 2001)). For example, a cell may have divided many times but still have a fully unmethylated tag. Another cell with tag methylation may have divided fewer times than a cell without tag methylation. Therefore, we purified phenotypically homogeneous cell groups and then sampled about eight tags per group. Average tag methylation was hypothesized to be proportional to average cell group mitotic age.

Leucocyte tag methylation during aging

As expected with stochastic replication errors (unprogrammed methylation and demethylation), leucocyte 5′ to 3′ tag methylation patterns were diverse (Fig 4), but several trends emerged. Average *NKX2-5* and *BGN* tag methylation were low in neutrophils and intermediate in B-cells (Fig 5). Age-related increases in neutrophil or B-cell methylation were not observed. By the logic of an epigenetic molecular clock, B-cells and neutrophils from individuals of all chronological ages have similar mitotic ages, but average B-cell mitotic ages are greater than those of neutrophils.

Immature T-cells differentiate into $CD3^+/4^+$ helper and $CD3^+/8^+$ suppressor T-cells in the thymus, and both types exhibited age-related methylation, with greater increases in $CD3^+/8^+$ suppressor T-cells (Fig 5). Therefore, consistent with long-lived cells that may commonly divide outside the bone marrow, average T-cell mitotic ages increase with chronological age. T-cell division is dependent on appropriate immunological stimulation, and therefore T-cell mitotic ages should vary within an individual. T-cells in older individuals were not uniformly more methylated, but had broad distributions with less methylated tags, and more methylated tags that were rarely observed in young individuals (Fig 5B).

A smaller number of *NKX2-5* tags were also examined from differentiating bone marrow nucleated red cell progenitors. An age-related increase in methylation was not observed for nucleated red cells, with an average *NKX2-5* tag methylation of 7·4%, which was significantly lower ($P = 0{\cdot}022$) than the average for neutrophils (Fig 5A). Overall, methylation patterns were consistent with average adult mitotic ages ordered as: nucleated red cells < neutrophils < B-cells < T-cells.

Tag methylation at birth

Newborns have the lowest chronological ages and potentially the lowest mitotic ages. However, unmethylated *BGN* and *NKX2-5* tags were present in leucocytes from young and old individuals, suggesting that cells with low mitotic ages are present in everyone (Fig 5B). There were no consistent methylation differences between newborns (<1 month old) and older individuals between 1 month and 5 years of age (data not shown).

To better characterize newborn mitotic ages, another tag, *SOX10*, was examined. *SOX10* appears to be a faster 'clock' because it becomes nearly fully methylated in hair by 2 years of age (Kim *et al*, 2006). *SOX10* methylation was significantly less in neutrophil, B-cells, and T-cells ($CD3^+/4^+$) from individuals <1 month old (Fig 6A and B), but subsequently nearly all *SOX10* tags were more methylated (>70%). Therefore, with respect to this faster epigenetic clock, leucocytes at birth have

© 2008 The Authors
Journal Compilation © 2008 Blackwell Publishing Ltd,

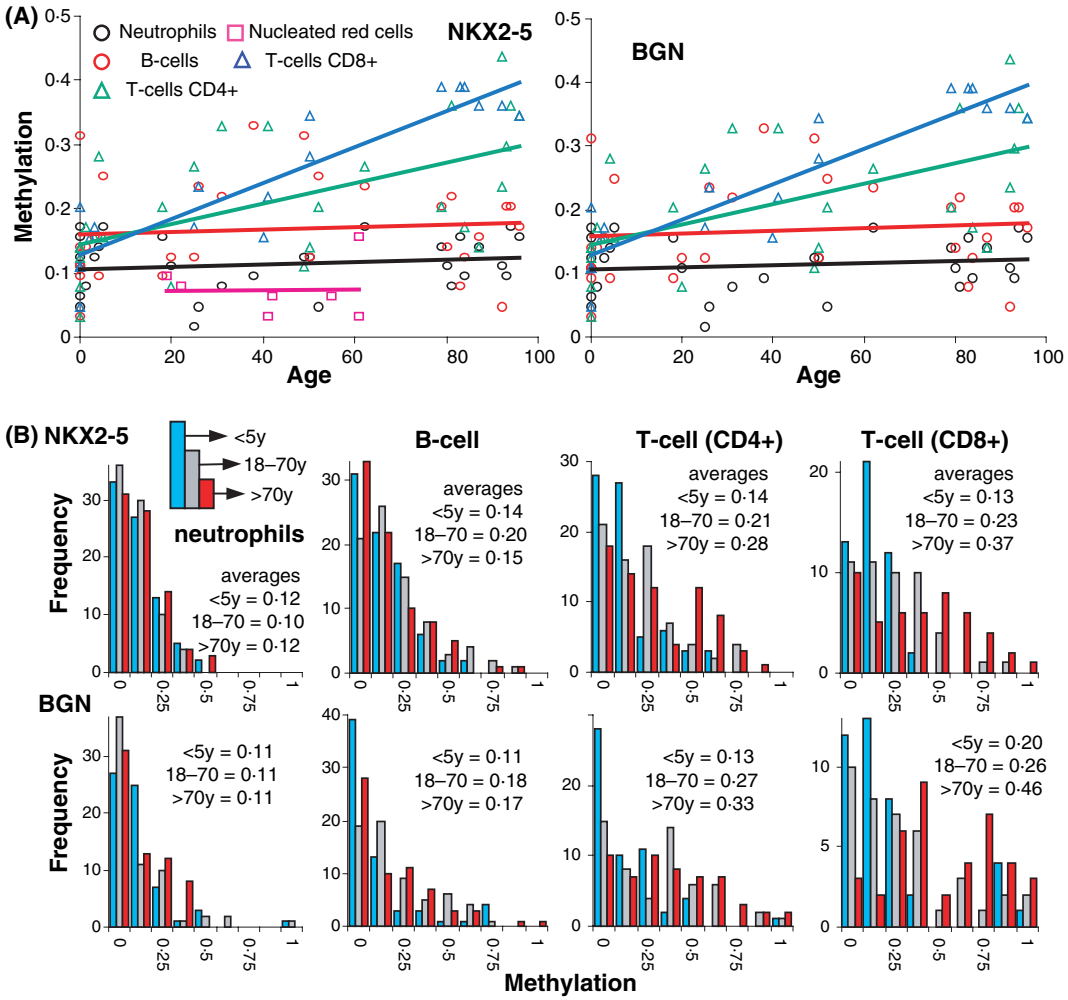

Fig 5. Tag methylation with aging. (A) Average *NKX2-5* or *BGN* (male individuals only) tag methylation with age for neutrophils (black circles and lines), B-cells (red circles and lines), CD4 T-cells (green triangles and lines), CD8 T-cells (blue triangles and lines), and nucleated red cells (purple squares and line). CD4 and CD8 T-cells exhibited age-related methylation with positive slopes (both $P < 0{\cdot}001$, linear mixed effects model). In contrast, neutrophils, B-cells and nucleated red cells did not exhibit age-related methylation, with trend lines not significantly different from a level line (all $P > 0{\cdot}10$, linear mixed effects model). (B) Frequency histograms of individual tag alleles, grouped by age (< 5 years old, blue bars; 18–70 years old, grey bars; >70 years old, red bars). There is a range of methylation for individual tags in every age group. The frequencies of more methylated T-cell tags increases with age, although unmethylated tags are still observed in the elderly.

lower mitotic ages than after 1 month of age. Tags ($n = 40$) from fetal haematopoietic tissues (three thymuses and two spleens, 20–32 weeks estimated gestation ages) also had low methylation with average *NKX2-5* methylation of 7·2% and *SOX10* methylation of 52%.

Tag methylation in CD34$^+$ cells

Although highly purified populations of HSCs are difficult to isolate because their numbers are few, HSCs and their more committed progenitors express a variety of surface markers including CD34. *NKX2-5* tags measured from CD34$^+$ bone marrow cells had low methylation levels similar to neutrophils (averages both 11%), with a small but non-significant age-related increase (Fig 7).

CD34$^+$ marrow cells are a heterogeneous population of cells with varying capabilities for haematopoiesis. A method to enumerate numbers of short-term haematopoietic progenitors is to culture colony-forming units (CFU), which are small colonies (several hundreds to thousands of cells) that arise from CD34$^+$ progenitors after 7–14 d and a limited number of divisions (Hao *et al*, 1998; Coulombel, 2004). The mitotic age of a CFU should be the age of its progenitors plus the relatively few divisions during culture. Recapitulating *in vivo* differentiation, *SOX10* cord blood CFU progeny methylation was similar to neonatal neutrophils (Fig 6C and Table I), and significantly lower than neutrophils from individuals >1 month old ($P = 0{\cdot}0015$) or adult CFU progeny ($P = 0{\cdot}014$). *SOX10* adult CFU progeny methylation was intermediate, significantly higher than cord blood CFU

© 2008 The Authors

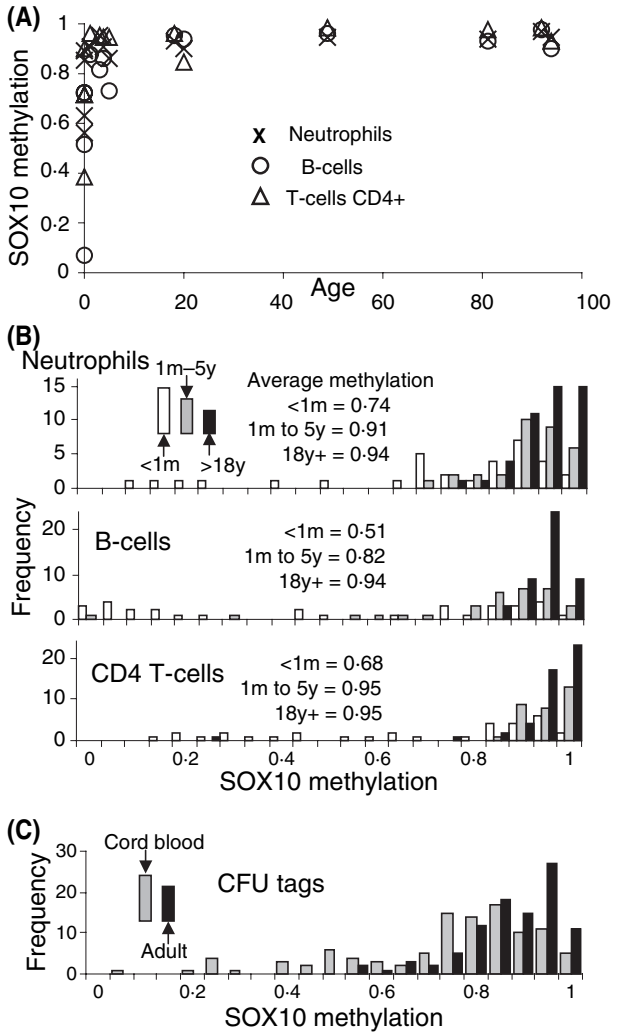

Fig 6. *SOX10* tag methylation. (A) Average *SOX10* tag methylation with age (neutrophils, crosses; B-cells, circles; CD4 T-cells, triangles). *SOX10* is a 'faster' molecular clock compared to *NKX2-5* and *BGN* because it becomes nearly fully methylated very early in life. (B) Frequency histograms of *SOX10* tag methylation. Newborn *SOX10* methylation (<1 month of age, white bars) was significantly less compared to older individuals (1 month to 5 years, grey bars; >18 years, black bars) for all cell types. (C) Frequency histogram of *SOX10* CFU progeny tags. Cord blood CFU progeny tags (grey) had a distribution similar to leucocytes from individuals <1 month old, and were also significantly less methylated than tags from individuals >1 month old. Adult CFU progeny tags (black) were significantly more methylated than cord blood CFU progeny tags, with a distribution more similar to adult neutrophils.

progeny but significantly lower ($P = 0{\cdot}008$) than neutrophils from individuals >1 month old. Similar to leucocytes, *NKX2-5* methylation was low in CFU progeny, with no significant difference between adult and cord blood sources. There were no obvious differences between methylation levels and CFU progeny phenotype. Specifically, lymphoid (B-cell) culture progeny had tag levels similar to circulating neutrophils, which suggest that the average higher tag methylation levels in circulating B-cells may be due to additional divisions from immunological stimulation after their release from the bone marrow.

Tag methylation in neoplasia

Tag methylation, and by inference the mitotic ages of normal leucocytes other than T-lymphocytes, was limited and did not significantly increase with age. Leukaemias are also thought to originate from HSCs or more committed progenitors, and transformation should allow for more divisions. Consistent with greater mitotic ages, blast *NKX2-5* tag methylation in six acute lymphoblastic (ALL) and 13 myeloid (AML) leukaemias was significantly greater than in normal neutrophils (Fig 8 and Table II). *BGN* tag methylation was also greater than in normal neutrophils, but the differences were not significant given the small numbers of leukaemias in males. There was a range of leukaemia tag methylation, with average values of some leukaemias similar to normal neutrophils. Leukaemia cell lines, which have presumably undergone many divisions *in vivo* and *in vitro*, had high (>70%) tag methylation levels (Table II).

Methylation in an expressed gene

To illustrate the relatively unique behaviour of 'neutral' clock tags, promoter methylation was also measured for *BMI1* (Table III), which is important for stem cell survival because haematopoiesis is deficient in mice lacking this gene (Park *et al*, 2003). *BMI1* expression is highest in HSCs and decreases with differentiation (Hosen *et al*, 2007). Consistent with expression in HSCs, *BMI1* promoter methylation was extremely low (<4%) in the progenitor $CD34^+$ population and in short-lived differentiated CFU, neutrophil, and red cell populations (Table III). *BMI1* methylation was higher in longer-lived B-cells (11%) and $CD4^+$ T-cells (46%). Although lymphoid cells have higher clock tag methylation, *BMI1* methylation did not simply appear to correlate with mitotic age because leukaemia cells have high mitotic ages but low *BMI1* methylation (2·3%), consistent with observations that *BMI1* is important for leukaemia survival (Lessard & Sauvageau, 2003).

Discussion

Haematopoiesis provides a well-defined test of a somatic molecular clock approach because its biology is well characterized (Shizuru *et al*, 2005) and many of its cell types can be purified by flow cytometry or culture. Tag data were generally consistent with the expected genealogies of human leucocytes, with T-lymphocytes but not B-lymphocytes or neutrophils demonstrating average age-related increases in methylation. Higher average tag methylation in T-cells from older individuals implies that these lineages may persist longer and divide more often relative to other leucocyte lineages.

© 2008 The Authors
Journal Compilation © 2008 Blackwell Publishing Ltd,

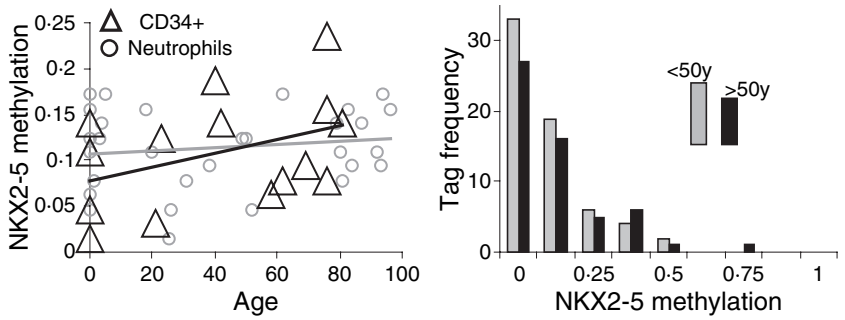

Fig 7. Average *NKX2-5* methylation of CD34 expressing bone marrow cells (black triangles and line) was similar to neutrophils (grey circles and line) but exhibited a non-significant ($P = 0{\cdot}14$) age-related increase in tag methylation. The histogram illustrates tag frequency from individuals less than (grey bars) and greater than (black bars) 50 years of age.

Table I. Average methylation of colony-forming unit (CFU) progeny *versus* peripheral blood neutrophils.

	SOX10 methylation		*NKX2-5* methylation	
Blood neutrophils				
<1 month old	74%* (4)		11% (6)	
>1 month old	93%* (10)		11% (24)	
Source	Cord blood	Adult	Cord blood	Adult
CFU-type				
BFU-E	68% (3)	86% (4)	10% (3)	4% (4)
CFU-GM	59% (3)	92% (4)	16% (3)	7% (4)
CFU-mix	80% (3)	84% (4)	11% (3)	7% (4)
Lymphoid	85% (4)	ND	10% (4)	ND
Average all CFUs	74% (13)*	87%* (12)	10% (13)	6% (12)

Parentheses indicate numbers of examined individuals or CFU colonies.

*Significant differences ($P < 0{\cdot}05$) are with *SOX10* methylation: blood neutrophil methylation is significantly lower in neonates (<1 month of age) or average adult CFU methylation, and average cord blood CFU methylation is significantly lower than adult CFU methylation.

BFU-E, erythroid burst-forming unit; CFU-GM, granulocyte-macrophage colony-forming unit; CFU-Mix, mixed lineage colony-forming unit.

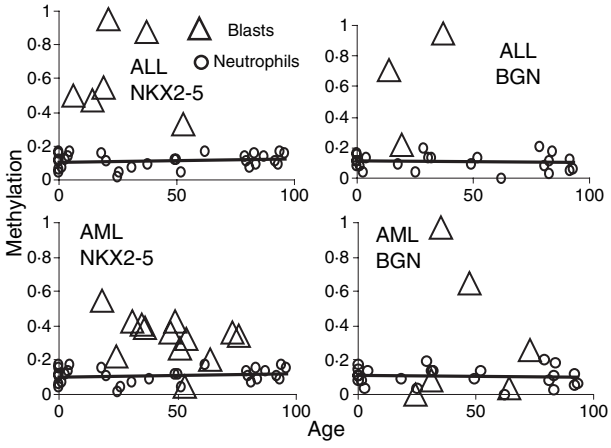

Fig 8. Average *NKX2-5* and *BGN* tag methylation values for ALL and AML leukaemic blasts (triangles) are scattered, with overall averages greater than normal neutrophils (circles and lines).

Table II. Average leukaemia *NKX2-5* and *BGN* tag methylation.

Sample type	*NKX2-5* methylation	*P* value*	*BGN* methylation	*P* value*
Neutrophils	11% (30)		11% (24)	
ALL†	62% (6)	0·0043	62% (3)	0·14
AML‡	36% (13)	<0·0001	34% (6)	0·22
Cell lines§	83% (5)	<0·001	90% (3)	0·0050

*Relative to neutrophils by a two-sided *t*-test.

†Pre-B-acute lymphoblastic leukaemia (ALL) ($n = 4$) and T-cell ALL ($n = 2$).

‡Acute myeloid leukaemia (AML) French-American-British (FAB) type M0 ($n = 1$), M2 ($n = 5$), M3 ($n = 4$), M4 ($n = 2$), M5 ($n = 1$).

§HL60, K562, Jurkat, Daudi, and Raji.

The number of samples are indicated in parentheses.

Table III. *BMI1* methylation.

Specimen type	*BMI1* methylation (%)	Specimens (# analysed tags)
$CD34^+$	3·2	9 (49)
CFU	2·1	9 (55)
Nucleated red cell	0·0	3 (17)
Neutrophil	1·3	6 (34)
B-Cell	11	6 (36)
T-Cell ($CD4^+$)	46	7 (42)
Leukaemia	2·3	6 (36)

HSC quiescence

The most pluripotent and undifferentiated haematopoietic progenitor is called the HSC, which is defined functionally by the ability to reconstitute haematopoiesis after transplantation. HSCs are few in number and cannot be purified to homogeneity because their cell surface markers overlap with other more differentiated progenitors (Purton & Scadden, 2007). Although HSC genomes were not directly measured in this study, it is possible to infer their mitotic ages by measuring the genomes of their more numerous and easy to isolate differentiated progeny. The genealogy of many cells can be divided into three sequential phenotypic phases (Fig 3B) – development from the zygote,

© 2008 The Authors

a stem cell phase, and differentiation (Shibata & Tavaré, 2006). Development and differentiation are programmed and therefore limited to specific times and numbers of divisions. Development typically occurs during the first few years of life and differentiation is usually limited to a few days or weeks. By contrast, the stem cell phase may vary because stem cells have the potential for limitless numbers of divisions. Given these restrictions, changes in the mitotic ages of differentiated cells largely reflect numbers of stem cell divisions (Fig 3B). Mature lymphocytes can also divide in response to appropriate immunological stimulation and therefore can increase their mitotic ages after their release from the bone marrow or thymus.

Neutrophil genealogy starts from the zygote, with a short period (~a year) of development to a HSC, a variable (~day to >100 years) HSC phase, and differentiation (~weeks) before death shortly after release from the bone marrow. More than a billion neutrophils are released everyday (Dancey *et al*, 1976). The inability to detect a significant age-related increase in clock tag methylation suggests that neutrophil genealogies are similar regardless of chronological age. New neutrophils in an old individual have essentially the same mitotic ages as new neutrophils in a younger individual. The lack of a significant age-related increase in neutrophil (or red cell progenitor and CD34 expressing cells) mitotic ages is consistent with the general idea that HSCs divide infrequently (Cheshier *et al*, 1999) and observations that leucocyte telomere lengths gradually shorten with aging. The rate of telomere length loss was consistent with less than one HSC division per year (Rufer *et al*, 1999; Brummendorf & Balabanov, 2006), which may not be detectable with the current relatively small amounts of clock tag data. Even in the colon, the increase in average *NKX2-5* tag methylation is only about 4% per decade (Fig 2).

It is uncertain how stem cells remain quiescent even though their non-stem cell progeny are mitotically active. Both hair growth and haematopoiesis exhibit high mitotic activity, but their stem cells appear to remain quiescent because significant age-related increases in the clock tag methylation of their differentiated progeny (hair follicle epithelium or neutrophils) were not observed (Fig 2). A common feature absent in glandular epithelium, such as the intestines and endometrium, is the cyclic physical destruction and reformation of the mitotic compartment. The hair stem cell compartment is the bulge, which is physically separated from the mitotic bulb compartment that contains the differentiated follicle cells (Fuchs *et al*, 2004). The bulb physically disappears every few years when the hair falls out, and a few months later bulge stem cells divide and their progeny migrate to reform the bulb. A similar cycle may occur with haematopoiesis because the functional HSC niche is adjacent to bone (Calvi *et al*, 2003; Zhang *et al*, 2003). Periodic niche disappearance could be balanced by the reformation of new haematopoietic niches when bone is remodelled (Aguila & Rowe, 2005).

HSC mitotic ages may be limited because they may only divide at the start of a niche repopulation cycle. Stem cell mitotic ages can also be limited by a related mechanism called clonal succession (Kay, 1965). In clonal succession, a pool of non-dividing stem cells successively produces differentiated cells. There appears to be a pool of HSCs because many normally circulate for short intervals in the blood (Wright *et al*, 2001). There may be sufficient numbers of HSCs, such that individual HSCs seldom occupy a marrow niche more than once during a lifetime. In this punctuated manner, the genealogy and mitotic ages of neutrophil or red cell progenitors may be similar regardless of chronological age because the same HSC does continuously produce progeny throughout life. HSC clonal succession has been inferred in cats (Abkowitz *et al*, 1990).

HSC development

Although mature HSCs are quiescent, a number of divisions must occur during their development from the zygote. Haematopoiesis starts in the first month after conception and serially moves from the yolk sac, to the liver, and finally to the bone marrow by birth (Zon, 1995). Residual haematopoiesis is often present in the liver at birth, but largely disappears by 1 month of age (Emery, 1956). At least some yolk sac haematopoietic cells appear to contribute to adult haematopoiesis (Samokhvalov *et al*, 2007), and fetal leucocytes could have lower mitotic ages because fetal HSC mitotic ages may increase as haematopoiesis migrates to the bone marrow. The *NKX2-5* and *BGN* clock tags detected no evidence of HSC development because their methylation levels were similar at all ages. However, a faster clock tag in *SOX10* was significantly less methylated in differentiated leucocytes from newborns and fetal tissues relative to tags from later ages (>1 month of age), suggesting that the mitotic ages of these cells are less than normal. From the logic of a molecular clock, HSC development or mitotic quiescence may not be fully established until shortly after birth.

Lymphocyte mitotic ages

Lymphocytes also differentiate from HSCs but are longer lived and divide when they encounter appropriate immunological stimuli. B-cell mitotic ages were significantly greater than neutrophils, which may reflect more divisions during differentiation or from subsequent antigen stimulation. Similar tag methylation levels between *in vitro* lymphoid cultures and neutrophil-containing CFUs suggests the additional methylation observed in circulating B-cells reflects subsequent antigen stimulation. Human B-cells appear to be produced throughout life (Nunez *et al*, 1996), and the lack of an age-related increase in errors (methylation) suggests a life-long balance between production and loss.

Helper and suppressor T-cells exhibited age-related increases in errors, and many T-cells in the elderly had high mitotic ages, consistent with numerous past immunological encounters. Long-lived T-cell lineages are not unexpected because T-cell differentiation requires a thymus, which physically regresses after puberty. T-cell numbers do not decrease with

© 2008 The Authors
Journal Compilation © 2008 Blackwell Publishing Ltd,

age (Linton & Dorshkind, 2004), and peripheral divisions, either homeostatic or from antigenic stimulation, may be required to maintain numbers after thymic regression. T-cells with low methylation levels were still present in the oldest individuals, consistent with the limited production of new T-cells throughout life (Douek *et al*, 1998; Linton & Dorshkind, 2004; Goronzy & Weyand, 2005), or that some T-cells seldom encounter immunological stimulation. Of note, most CpG sites do not appear to exhibit age-related methylation changes in human T-cells (Tra *et al*, 2002). However, the current sequencing approach may allow for more precise methylation measurements relative to hybridization approaches. Age-related telomere shortening was also observed in T-cells, with greater telomere shortening in T-cells relative to neutrophils (Rufer *et al*, 1999).

Neoplastic leucocyte mitotic ages

Another mechanism that increases mitotic age is transformation, which removes limitations on cell divisions. Consistent with unregulated mitotic activity, average ALL and AML blast mitotic ages were greater than neutrophils. There was a range of leukaemia mitotic ages, suggesting different leukaemias accumulate different numbers of divisions before or after transformation. Leukaemia, like normal haematopoiesis, appears to have a stem cell hierarchy (Bonnet & Dick, 1997; Warner *et al*, 2004), and leukaemia blast methylation may reflect the numbers of errors in a leukaemic stem cell (LSC). The genealogy of leukaemia also starts with the zygote, with intervening HSC, LSC, and blast phenotypes.

Limitations and caveats

Defining the ancestry of a genome can be problematic because genomes constantly accumulate random variations. There is not a single human genome but rather many human genomes that vary within and between individuals. Although some variations confer phenotypes, most changes appear to represent neutral polymorphisms that arise from random replication errors (Kidd *et al*, 2004). Such variations support the idea that life evolves through unprogrammed random changes and selection, rather than some sort of 'design'.

Similarly, it may be hard to define the epigenetic state or epigenome of a given cell type because variations are likely unless development and differentiation are somehow perfectly programmed. Most epigenetic variations are unlikely to modify cell phenotype because only a fraction of genes are expressed and minor changes may not significantly change expression. The heterogeneity of methylation patterns in homogenous leucocyte populations observed in this study and many others (see for example Zhu *et al*, 1999) is evidence that random unprogrammed epigenetic variation also occurs during development, differentiation, and aging.

The relatively small amounts of data presented in this study should be considered preliminary evidence that certain leucocyte methylation patterns appear predominantly to represent unprogrammed replication errors. Additional studies with a variety of different cell types (Fig 2) are also consistent with this hypothesis. Certain types of cell lineages appear to divide throughout life whereas the mitotic ages of other cell types are more limited. Even this vague outline of the past is valuable because direct measurements of human somatic cell mitotic ages are impractical. Future challenges include the identification of additional informative tags and the development of more formal quantitative approaches to analyze epigenetic variations. Molecular phylogenetic methods have revolutionized the study of species evolution because it is easy to measure and compare variation between present-day genomes. It is possible that the ancestries of all cells, including somatic cells, are recorded within their genomes.

Acknowledgements

This work was supported by grants from the National Institutes of Health (CA111940) and the Norris Cancer Center. ST is a Royal Society-Wolfson Research Merit Award holder, and is supported in part by a grant from Cancer Research UK. We thank Jo Ann Woo, Nancy Quash, and Carrie Eckstram for technical assistance, and Sylvia Chavira for her expertise with flow cytometry.

References

Abkowitz, J.L., Linenberger, M.L., Newton, M.A., Shelton, G.H., Ott, R.L. & Guttorp, P. (1990) Evidence for the maintenance of hematopoiesis in a large animal by the sequential activation of stem-cell clones. *Proceedings of the National Academy of Sciences of the United States of America*, **87,** 9062–9066.

Aguila, H.L. & Rowe, D.W. (2005) Skeletal development, bone remodeling, and hematopoiesis. *Immunological Reviews*, **208,** 7–18.

Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes and Development*, **16,** 6–21.

Bonnet, D. & Dick, J.E. (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Medicine*, **3,** 730–737.

Brero, A., Leonhardt, H. & Cardoso, M.C. (2006) Replication and translation of epigenetic information. *Current Topics in Microbiology and Immunology*, **301,** 21–44.

Bromham, L. & Penny, D. (2003) The modern molecular clock. *Nature Review Genetics*, **4,** 216–224.

Brummendorf, T.H. & Balabanov, S. (2006) Telomere length dynamics in normal hematopoiesis and in disease states characterized by increased stem cell turnover. *Leukemia*, **20,** 1706–1716.

Calvi, L.M., Adams, G.B., Weibrecht, K.W., Weber, J.M., Olson, D.P., Knight, M.C., Martin, R.P., Schipani, E., Divieti, P., Bringhurst, F.R., Milner, L.A., Kronenberg, H.M. & Scadden, D.T. (2003) Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature*, **425,** 841–846.

Cheshier, S.H., Morrison, S.J., Liao, X. & Weissman, I.L. (1999) *In vivo* proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, **96,** 3120–3125.

© 2008 The Authors

Chu, M.W., Siegmund, K.D., Eckstam, C.L., Kim, J.Y., Yang, A.S., Kanel, G.C., Tavaré, S. & Shibata, D. (2007) Lack of increases in methylation at three CpG-rich genomic loci in non-mitotic adult tissues during aging. *BMC Medical Genetics*, **8,** 50.

Coulombel, L. (2004) Identification of hematopoietic stem/progenitor cells: strength and drawbacks of functional assays. *Oncogene*, **23,** 7210–7222.

Dancey, J.T., Deubelbeiss, K.A., Harker, L.A. & Finch, C.A. (1976) Neutrophil kinetics in man. *Journal of Clinical Investigation*, **58,** 705–715.

Douek, D.C., McFarland, R.D., Keiser, P.H., Gage, E.A., Massey, J.M., Haynes, B.F., Polis, M.A., Haase, A.T., Feinberg, M.B., Sullivan, J.L., Jamieson, B.D., Zack, J.A., Picker, L.J. & Koup, R.A. (1998) Changes in thymic function with age and during the treatment of HIV infection. *Nature*, **396,** 690–695.

Emery, J.L. (1956) The distribution of haemopoietic foci in the infantile human liver. *Journal of Anatomy*, **90,** 293–297.

Fuchs, E., Tumbar, T. & Guasch, G. (2004) Socializing with the neighbors: stem cells and their niche. *Cell*, **116,** 769–778.

Goronzy, J.J. & Weyand, C.M. (2005) T cell development and receptor diversity during aging. *Current Opinion in Immunology*, **17,** 468–475.

Hao, Q.L., Smogorzewska, E.M., Barsky, L.W. & Crooks, G.M. (1998) *In vitro* identification of single $CD34^{+}CD38$- cells with both lymphoid and myeloid potential. *Blood*, **91,** 4145–4151.

Hedges, S.B. & Kumar, S. (2003) Genomic clocks and evolutionary timescales. *Trends in Genetics*, **19,** 200–206.

Holliday, R. (1987) The inheritance of epigenetic defects. *Science*, **238,** 163–170.

Hosen, N., Yamane, T., Muijtjens, M., Pham, K., Clarke, M.F. & Weissman, I.L. (2007) Bmi-1-green fluorescent protein-knock-in mice reveal the dynamic regulation of bmi-1 expression in normal and leukemic hematopoietic cells. *Stem Cells*, **25,** 1635–1644.

Issa, J.P. (2000) CpG-island methylation in aging and cancer. *Current Topics in Microbiology and Immunology*, **249,** 101–118.

Jones, P.A. & Laird, P.W. (1999) Cancer epigenetics comes of age. *Nature Genetics*, **21,** 163–167.

Kay, H.E. (1965) How many cell-generations? *Lancet*, **15,** 418–419.

Kidd, K.K., Pakstis, A.J., Speed, W.C. & Kidd, J.R. (2004) Understanding human DNA sequence variation. *Journal of Heredity*, **95,** 406–420.

Kim, J.Y., Siegmund, K.D., Tavaré, S. & Shibata, D. (2005a) Age-related human small intestine methylation: evidence for stem cell niches. *BMC Medicine*, **3,** 10.

Kim, J.Y., Tavaré, S. & Shibata, D. (2005b) Counting human somatic cell replications: methylation mirrors endometrial stem cell divisions. *Proceedings of the National Academy of Sciences of the United States of America*, **102,** 17739–17744.

Kim, J.Y., Tavaré, S. & Shibata, D. (2006) Human hair genealogies and stem cell latency. *BMC Biology*, **4,** 2.

Lessard, J. & Sauvageau, G. (2003) Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. *Nature*, **423,** 255–260.

Linton, P.J. & Dorshkind, K. (2004) Age-related changes in lymphocyte development and function. *Nature Immunology*, **5,** 133–139.

Morgan, H.D., Santos, F., Green, K., Dean, W. & Reik, W. (2005) Epigenetic reprogramming in mammals. *Human Molecular Genetics*, **14,** R47–R58.

Nunez, C., Nishimoto, N., Gartland, G.L., Billips, L.G., Burrows, P.D., Kubagawa, H. & Cooper, M.D. (1996) B cells are generated throughout life in humans. *Journal of Immunology*, **156,** 866–872.

Park, I.K., Qian, D., Kiel, M., Becker, M.W., Pihalja, M., Weissman, I.L., Morrison, S.J. & Clarke, M.F. (2003) Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature*, **423,** 302–305.

Purton, L.E. & Scadden, D.T. (2007) Limiting factors in murine hematopoietic stem cell assays. *Cell Stem Cell*, **1,** 263–270.

Rufer, N., Brummendorf, T.H., Kolvraa, S., Bischoff, C., Christensen, K., Wadsworth, L., Schulzer, M. & Lansdorp, P.M. (1999) Telomere fluorescence measurements in granulocytes and T lymphocyte subsets point to a high turnover of hematopoietic stem cells and memory T cells in early childhood. *Journal of Experimental Medicine*, **190,** 157–167.

Samokhvalov, I.M., Samokhvalova, N.I. & Nishikawa, S. (2007) Cell tracing shows the contribution of the yolk sac to adult haematopoiesis. *Nature*, **446,** 1056–1061.

Shibata, D. & Tavaré, S. (2006) Counting divisions in a human somatic cell tree: how, what and why? *Cell Cycle*, **5,** 610–614.

Shizuru, J.A., Negrin, R.S. & Weissman, I.L. (2005) Hematopoietic stem and progenitor cells: clinical and preclinical regeneration of the hematolymphoid system. *Annual Review of Medicine*, **56,** 509–538.

Sjoblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., Mandelker, D., Leary, R.J., Ptak, J., Silliman, N., Szabo, S., Buckhaults, P., Farrell, C., Meeh, P., Markowitz, S.D., Willis, J., Dawson, D., Willson, J.K., Gazdar, A.F., Hartigan, J., Wu, L., Liu, C., Parmigiani, G., Park, B.H., Bachman, K.E., Papadopoulos, N., Vogelstein, B., Kinzler, K.W. & Velculescu, V.E. (2006) The consensus coding sequences of human breast and colorectal cancers. *Science*, **314,** 268–274.

Sprent, J. & Tough, D.F. (1994) Lymphocyte life-span and memory. *Science*, **265,** 1395–1400.

Tra, J., Kondo, T., Lu, Q., Kuick, R., Hanash, S. & Richardson, B. (2002) Infrequent occurrence of age-dependent changes in CpG island methylation as detected by restriction landmark genome scanning. *Mechanisms of Ageing and Development*, **123,** 1487–1503.

Velicescu, M., Weisenberger, D.J., Gonzales, F.A., Tsai, Y.C., Nguyen, C.T. & Jones, P.A. (2002) Cell division is required for de novo methylation of CpG islands in bladder cancer cells. *Cancer Research*, **62,** 2378–2384.

Warner, J.K., Wang, J.C., Hope, K.J., Jin, L. & Dick, J.E. (2004) Concepts of human leukemic development. *Oncogene*, **23,** 7164–7177.

Wright, D.E., Wagers, A.J., Gulati, A.P., Johnson, F.L. & Weissman, I.L. (2001) Physiological migration of hematopoietic stem and progenitor cells. *Science*, **294,** 1933–1936.

Yatabe, Y., Tavaré, S. & Shibata, D. (2001) Investigating stem cells in human colon by using methylation patterns. *Proceedings of the National Academy of Sciences of the United States of America*, **98,** 10839–10844.

Zhang, J., Niu, C., Ye, L., Huang, H., He, X., Tong, W.G., Ross, J., Haug, J., Johnson, T., Feng, J.Q., Harris, S., Wiedemann, L.M., Mishina, Y. & Li, L. (2003) Identification of the haematopoietic stem cell niche and control of the niche size. *Nature*, **425,** 836–841.

Zhu, X., Deng, C., Kuick, R., Yung, R., Lamb, B., Neel, J.V., Richardson, B. & Hanash, S. (1999) Analysis of human peripheral blood T cells and single-cell-derived T cell clones uncovers extensive clonal CpG island methylation heterogeneity throughout the genome. *Proceedings of the National Academy of Sciences of the United States of America*, **96,** 8058–8063.

Zon, L.I. (1995) Developmental biology of hematopoiesis. *Blood*, **86,** 2876–2891.

© 2008 The Authors
Journal Compilation © 2008 Blackwell Publishing Ltd,