



## Supplementary Materials for

### **Variation in cancer incidence among tissues can be explained by the number of stem cell divisions**

Cristian Tomasetti\* and Bert Vogelstein\*

\*Corresponding author. E-mail: [ctomasetti@jhu.edu](mailto:ctomasetti@jhu.edu) (C.T.); [vogelbe@jhmi.edu](mailto:vogelbe@jhmi.edu) (B.V.)

Published 2 January 2015, *Science* **347**, 78 (2015)  
DOI: 10.1126/science.1260825

#### **This PDF file includes:**

Materials and Methods  
Fig. S1  
Table S1  
References

## Materials and Methods

### Overview

The parameters listed in Table S1 correspond to the average size of the relevant normal tissues and were obtained from the literature directly or calculated using values from the literature. The references are provided below in the sections describing each cancer type. When available, some tissues estimates have been given for both specific anatomic locations as well as for the overall tissue (e.g. osteosarcoma). Lifetime incidences are all based on U.S. data. In general, the data are from [seer.cancer.gov](http://seer.cancer.gov) (<http://seer.cancer.gov/statfacts/> and [http://seer.cancer.gov/archive/csr/1975\\_2010/results\\_merged/topic\\_lifetime\\_risk.pdf](http://seer.cancer.gov/archive/csr/1975_2010/results_merged/topic_lifetime_risk.pdf)). If the data were not available there, we used [cdc.gov](http://cdc.gov), [cancer.org](http://cancer.org), or individual publications addressing a specific (typically rare) type of cancer. Because the confidence intervals for the stem cell parameter estimates in the literature were not provided in all cases, we performed a sensitivity analysis assuming a 100-fold variation (plus and minus) for the estimates (see Statistical Analysis section below). We assessed a total of 31 different cancer types. Other cancer types could not be assessed, largely because details about the normal stem cells maintaining the tissue in homeostasis have not yet been agreed upon or accurately quantified.

### Acute myeloid leukemia

The lifetime incidence of acute myeloid leukemia is 0.41% ([www.seer.cancer.gov](http://www.seer.cancer.gov)) (3). There are a total of  $\sim 3 \cdot 10^{12}$  blood cells and  $7.5 \cdot 10^{11}$  nucleated cells in the bone marrow (32). 0.18% of these normal bone marrow cells are CD34+ (33). The CD34+ are a heterogeneous population of cells and only about 10% of the CD34+ are hematopoietic stem cells (HSC), phenotypically defined as (Lin-CD34+CD38-CD90+CD45RA-) (34). The number of hematopoietic stem cells therefore equals  $7.5 \cdot 10^{11} \cdot 0.0018 \cdot 0.1 = 1.35 \cdot 10^8$ . Hematopoietic stem cells have been estimated to divide every  $\sim 15$  (35), 17 (36), 30 (37, 38), 57 (39) days. Thus, we assume they divide every 30 days.

### Basal cell carcinoma

The lifetime incidence of cutaneous basal cell carcinoma has been estimated to be  $\sim 30\%$  (40, 41), making it the most common form of cancer (42, 43). The total number of cells in the epidermis, the outermost layer of the skin, is  $\sim 1.8 \cdot 10^{11}$  (32). Cutaneous basal cell carcinomas originate from long-term resident progenitor cells present in the basal (innermost) layer of the epidermis (44). These basal cells possess proliferative potential, while the suprabasal layers are composed of terminally differentiating keratinocytes (45), (25). The basal layer is found both around the hair follicles as well as in the interfollicular epidermis (25). The number of basal cells has been estimated to be  $\sim 3 \cdot 10^4$  cells/mm<sup>2</sup> (46) and 10% of these cells are believed to be stem cells (47, 48). The median height of a US woman is  $\sim 5' 4.5''$ , and that of a US man is  $\sim 5' 9.5''$ , while the median weight is  $\sim 160$  lbs and  $\sim 190$  lbs, respectively ([www.census.gov](http://www.census.gov)). Considering both sexes together, the average surface area of the skin is  $\sim 1.94$  m<sup>2</sup> (49). The median number of basal cells in the epidermis is therefore  $\sim 3 \cdot 10^4 \cdot 1.94 \cdot 10^6 = 5.82 \cdot 10^{10}$ , and the number of stem cells

is  $\sim 5.82 \cdot 10^9$ . Epidermal stem cells are estimated to divide every 48 days (50, 51).

### Chronic lymphocytic leukemia

The lifetime incidence of chronic lymphocytic leukemia is 0.52% (www.seer.cancer.gov) (3). There are a total of  $\sim 3 \cdot 10^{12}$  blood cells and  $7.5 \cdot 10^{11}$  nucleated cells in the bone marrow (32). 0.18% of these normal bone marrow cells are CD34+ (33). The CD34+ are a heterogeneous population of cells and only about 10% of the CD34+ are hematopoietic stem cells, phenotypically defined as (Lin-CD34+CD38-CD90+CD45RA-) (34). The number of hematopoietic stem cells therefore equals  $7.5 \cdot 10^{11} \cdot 0.0018 \cdot 0.1 = 1.35 \cdot 10^8$ . Hematopoietic stem cells divide every  $\sim 30$  days (see acute myeloid leukemia for details).

### Colorectal adenocarcinoma

The lifetime incidence of colorectal cancer is  $\sim 4.8\%$  (www.seer.cancer.gov) (3). Almost all colorectal cancers are adenocarcinomas. The large intestine is on average  $\sim 1.6$  meters long and 5-8 cm in diameter, resulting in an internal surface area of approximately  $3,300 \text{ cm}^2$ . There are a total of  $\sim 1.5 \cdot 10^7$  crypts ( $14,000 \text{ crypts/cm}^2$ ) in the large intestine, each of which contains  $\sim 2,000$  cells (52-56). It has been estimated that each colonic crypt contains only 15-20 stem cells (57) or even fewer (55), with a turnover of approximately 3-7 days (55, 58-60). Thus, we estimate that the total number of stem cells in the large intestine is  $\sim 1.5 \cdot 10^7 \cdot 15 \sim 2 \cdot 10^8$  and that stem cells divide, on average, every 5 days.

### Colorectal adenocarcinoma in FAP patients

Familial adenomatous polyposis (FAP) is an inherited condition caused by defects in the APC genes. Virtually all patients with FAP develop colorectal adenocarcinomas unless a colectomy has been performed, so their lifetime incidence is 100%. All other parameters are the same as for sporadic colorectal adenocarcinomas.

### Colorectal adenocarcinoma in patients with HNPCC

HNPCC (Hereditary Non-polyposis Colorectal Cancer, also called Lynch Syndrome) accounts for 1-3% of colorectal cancers, and the lifetime incidence of colorectal cancer for those with HNPCC has been estimated to be  $\sim 50\%$  (61), (62). All other parameters are the same as for sporadic colorectal adenocarcinomas.

### Duodenal adenocarcinoma

Forty to fifty per cent of the small intestinal adenocarcinomas occur in the duodenum (63). Thus, the lifetime incidence of duodenal adenocarcinoma is  $\sim 0.002 \cdot 0.35 \cdot 0.45 = 0.0003$ . The duodenum is approximately 25 cm long and  $\sim 2.5$  cm in diameter, resulting in a total surface of approximately  $200 \text{ cm}^2$ , representing  $\sim 4\%$  of the surface of the small intestine. Based on the numbers provided for the small intestine, the total number of cells in the duodenum can be calculated to be  $\sim 0.04 \cdot 1.7 \cdot 10^{10} = 7 \cdot 10^8$  and the number of stem cells is  $\sim 0.04 \cdot 1 \cdot 10^8 = 4 \cdot 10^6$ . Also, as noted in the section on the small intestine, the duodenal stem cells divide  $\sim 1/3$  as frequently as those of the colon (64, 65).

### Duodenal adenocarcinoma in FAP patients

Familial adenomatous polyposis (FAP) is an inherited condition caused by defects in the APC gene. Those with FAP have a lifetime incidence of duodenum carcinoma of 3-4% (40, 41). All other parameters are the same as those for sporadic duodenal adenocarcinomas.

#### Esophageal squamous cell carcinoma

The lifetime incidence of cancer of the esophagus is ~0.51% (3). The vast majority of these cancers are either adenocarcinomas or squamous cell carcinomas, and the ratio of these two types has changed considerably in the last few decades. Currently, ~38% of esophageal cancers are squamous cell carcinomas (66, 67). The lifetime incidence of esophageal squamous cell carcinoma is therefore  $0.0051 \cdot 0.38 = 0.001938$ . The esophagus is 18 to 26 cm long and 2 to 3 cm in diameter (68), resulting in an average surface area of  $\sim 172.79 \text{ cm}^2$ . The area of a squamous cell in the basal layer is  $\sim 80 \mu\text{m}^2$  (69). The fraction of stem cells in the basal layer has been estimated to be 0.4% of the total basal layer cells (70), so there are a total of  $\sim 2.16 \cdot 10^8$  basal cells and  $8.64 \cdot 10^5$  stem cells in the basal layer. As the normal mucosal epithelium is  $\sim 10$ -20 layers thick, the total number of epithelial cells in the esophagus is estimated to be  $\sim 2.16 \cdot 10^8 \cdot 15 = 3.24 \cdot 10^9$ . Cells have been estimated to turn over  $\sim$  every 21 days (71, 72).

#### Gallbladder non-papillary adenocarcinoma

There will be an estimated 10,650 new cases of gallbladder cancer in 2014 in the U.S., versus 136,830 new cases of colorectal cancer, and because the lifetime risk of colorectal cancer is 4.8%, we estimate the lifetime incidence of gallbladder cancer as  $\sim (10650/136830) \cdot 0.048 = 0.003736$  (3). About 75% of all gallbladder cancers are non papillary adenocarcinomas (73). Thus, the lifetime incidence of gallbladder non papillary adenocarcinoma is  $0.003736 \cdot 0.75 = 0.0028$ . There are  $\sim 1.6 \cdot 10^8$  epithelial cells in the gallbladder (32) and about 1% of them exhibit stem cell properties (74). Tissue renewal is very slow, with stem cell division once every  $\sim 625$  days (75).

#### Glioblastoma multiforme

The lifetime incidence of brain cancer is  $\sim 0.6\%$  (www.seer.cancer.gov) (3). Gliomas represent 81% of brain cancers, with glioblastomas being the most common among them (45%) (76). Thus, the lifetime incidence of glioblastoma is  $0.006 \cdot 0.81 \cdot 0.45 = 0.00219$ . The cell of origin of glioblastomas is the astrocyte, a type of glial cell. While it was previously believed that there are  $>10$  times more glial cells than neurons in the human brain (see for example (32) and the textbook references cited therein), recent evidence indicates that the ratio is much closer to 1:1, with an estimated  $84.6 \cdot 10^9$  glial cells and  $86.1 \cdot 10^9$  neurons in the human brain (77, 78). Because the great majority of glial cells are astrocytes (79), we estimate that there are  $84.6 \cdot 10^9$  astrocytes in the brain. Astrocytes are derived from neural stem cells (79), whose proportion has been estimated to be  $\sim 0.16\%$  (80). Thus, we estimate that the number of neural stem cells among the astrocytes is equal to  $84.6 \cdot 10^9 \cdot 0.0016 = 1.35 \cdot 10^8$ , with virtually no cell division after birth (81).

#### Head and neck squamous cell carcinoma

The lifetime incidence of oral or pharyngeal cancer is 1.09% and is 0.36% for laryngeal cancer, and the vast majority of these are squamous cell carcinomas (www.seer.cancer.gov) (3). The lifetime incidence of all head and neck squamous cell carcinomas, which include oral, pharyngeal, and laryngeal subtypes, is therefore ~1.45%. *Human Papillomavirus* (HPV) infection is an established contributor to oral and pharyngeal squamous cell carcinomas (82-85), and recent evidence of its role in laryngeal cancers has been published (86). The overall prevalence of oral HPV-16 infections in the general US population is estimated to be 1% (87). The odds ratio for head and neck cancer in those that are HPV-16 positive has been estimated to be 5.75 (88). Given the rarity of the disease, we can approximate the incidence ratio by the odds ratio. As  $0.0145 = 0.01 \cdot 5.75 \cdot x + 0.99 \cdot x$  when  $x \approx 0.0138$ , the lifetime incidence of head and neck cancer is 0.07935 for those infected with HPV-16, and 0.0138 for those not infected. The average total surface of the oral mucosa is ~172 to 197 cm<sup>2</sup> (89, 90), while the pharyngeal surface area is  $\sim 6.085 \cdot 2.56 \cdot \pi \sim 49$  cm<sup>2</sup> (see Fig. 2 and Table 1 of (91)), and the total surface of the larynx (the laryngeal lumen, which is tubular in shape and lined with mucosa) is  $\sim (6.31+5.12)/2 \cdot \pi \cdot ((4.26+3.7)/2 + (3.39+2.71)/2)/2 \sim 63$  cm<sup>2</sup> (see Table 2 in (92)), for a total surface area of the oral cavity, pharynx and larynx of ~296.5 cm<sup>2</sup>. The area of a cell in the basal layer of the oral cavity epithelium is ~80μm<sup>2</sup> (69). The fraction of stem cells in the basal layer is 3-7% (93), and given that there are a total of  $\sim 3.71 \cdot 10^8$  basal cells,  $1.85 \cdot 10^7$  stem cells are found in the basal layer. As the epithelium of the buccal mucosa is 40-50 layers thick (94), the total number of epithelial cells in the head and neck mucosa is  $\sim 3.71 \cdot 10^8 \cdot 45 = 1.67 \cdot 10^{10}$ . Stem cells in the oral mucosa divide every 14 to 20 days (39, 65).

#### Hepatocellular carcinoma

The lifetime incidence of liver and intrahepatic bile duct cancer is ~0.86% (www.seer.cancer.gov) (3), and ~90% of these cancers are hepatocellular carcinomas (HCC)(95). About 10% of all HCC are associated with *Hepatitis C* infection (HCV) (95, 96), and ~1% of the US population is infected with HVC (97). Thus, the incidence ratio for those with HCV infection is 0.1/.01, i.e., ten times greater than those without HCV infection. Based on the calculation that  $0.0086 \cdot 0.9 = 0.01 \cdot 10 \cdot x + 0.99 \cdot x$  when  $x \approx 0.0071$ , the lifetime incidence of HCC is ~0.071 for those infected with HCV, and ~0.0071 for individuals not infected with HCV. There are  $\sim 2.41 \cdot 10^{11}$  hepatocytes in the liver (32). Hepatic stem cells' (hHpSC) represent 0.5-2% of the total hepatocytes (98). Cell turnover has been estimated to be very low, with estimates ranging from 300 to 500 days (81).

#### Lung adenocarcinoma

The lifetime incidence of lung and bronchus cancer is 6.9% (3). Approximately 87.5% of lung cancers are non-small cell lung cancers (40% adenocarcinomas, 30% squamous cell carcinomas, 9% large cell carcinomas) while the remaining 20% are small cell lung cancers (99). Men who smoke are ~23 times more likely to develop lung cancer, and women who smoke are ~13 times more likely to develop lung cancer than nonsmokers (5). Thus, combining both sexes, the lifetime incidence is 18 times higher in smokers than nonsmokers. Smoking prevalence rose from 42.4% in 1965 to 18.1% in 2012 (100). Assuming an average smoking prevalence of 30%, and because  $0.069 \cdot 0.4 = 0.3 \cdot 18 \cdot x +$

$0.7 \cdot x$  when  $x \approx 0.0045$ , it follows that the lifetime incidence of lung adenocarcinoma is approximately 0.0045 for nonsmokers and 0.081 for smokers. Lungs and bronchus have a total of approximately  $43.4 \cdot 10^{10}$  cells, with the vast majority being endothelial, interstitial and alveolar cells (32). Lung adenocarcinomas typically originate in the peripheral lung tissue. Bronchio-alveolar stem cells have recently been identified (101-103), and while their precise characterization is still a matter of debate, the estimates for their frequency are rather similar, ranging from 0.4% of all lung cells (101) (for a total of  $0.004 \cdot 43.4 \cdot 10^{10} \sim 1.74 \cdot 10^9$  stem cells), to 1% of all alveolar cells type II (103) (for a total of  $0.7 \cdot 10^9$  stem cells, because there are  $\sim 7 \cdot 10^{10}$  alveolar type II cells in the two lungs of a human (32)). We take the average of these values as our estimate of the number of stem cells, i.e.,  $1.22 \cdot 10^9$ . Finally, the stem cell division rate for lung tissue (specifically the alveoli) has been estimated to be 7% per year (103).

### Medulloblastoma

The lifetime incidence of brain cancer is 0.6% (www.seer.cancer.gov) (3). Between 2005 and 2009, in the United States, the ratio between the number of newly diagnosed medulloblastomas and glioblastomas, across all ages, was  $\sim 0.05$  (104). A similar proportion (4.3%) was found through the analysis of age-specific incidence rates provided by the Central Brain Tumor Registry of the United States (www.cbtrus.org). Thus, given that the lifetime incidence of glioblastoma is 0.00219 (see above), the lifetime incidence of medulloblastoma is  $0.00219 \cdot 0.05 = 0.00011$ . Medulloblastomas form in the cerebellum. It is currently thought that medulloblastomas' cell of origin is the neural stem cell, which gives rise to both neurons and glial cells, with neuronal differentiation far more common (105). It has been estimated that the cerebellum contains  $69 \cdot 10^9$  neurons and  $16 \cdot 10^9$  glial cells (77). Neural stem cells have been estimated to represent  $\sim 0.16\%$  of the total cerebellar neuronal and glial cells (80). Thus, we estimate that the number of neural stem cells in the cerebellum is  $(69+16) \cdot 10^9 \cdot 0.0016 = 1.36 \cdot 10^8$ , and these cells do not divide after birth (81).

### Medullary thyroid carcinoma

The lifetime incidence of thyroid cancer is  $\sim 1.08\%$  (www.seer.cancer.gov) (3). Only  $\sim 3\%$  of these cancers are derived from parafollicular cells (C cells) (106-108), which are the cell of origin for medullary carcinomas. Thus, the lifetime incidence of medullary thyroid carcinomas is  $\sim 0.0108 \cdot 0.03 = 0.000324$ . There are approximately 10 follicular cells in the thyroid for every parafollicular cell (109, 110). As there are  $\sim 10^{10}$  follicular cells in the thyroid (32), there should be  $\sim 10^9$  parafollicular cells in the thyroid. Both follicular and parafollicular cells are derived from the pharyngeal endoderm (111), and recent evidence indicates a possible common stem cell of origin (112, 113). The stem cell population has been estimated to represent  $\sim 0.1-1.4\%$  of the total thyroid cells (112, 114-116). Thus we assume there are  $0.0065 \cdot 10^9 = 6.5 \cdot 10^6$  stem cells that can give rise to medullary thyroid carcinomas. Cell turnover has been estimated to be very low, with stem cell divisions occurring once every 8.5 – 14.4 years (116, 117).

### Melanoma

The lifetime incidence of melanoma is 2.03% (www.seer.cancer.gov) (3). Melanoma originates from melanocytes, cells residing in the stratum basale (basal layer) of the

epidermis (118). The average number of melanocytes has been estimated to be  $3.8 \cdot 10^9$  (32). Other estimates suggests that there are  $\sim 1.25 \cdot 10^3$  cells/mm<sup>2</sup> (119, 120). As individuals have a median surface area of  $\sim 1.94$  m<sup>2</sup> (see basal cell carcinoma section), the median number of melanocytes in the epidermis would be  $\sim 1.25 \cdot 10^3 \cdot 1.94 \cdot 10^6 = 2.425 \cdot 10^9$ , an estimate close to that of (32). This estimate is also consistent with studies that show melanocytes represent 5-10% of the cells in the basal layer (121), which we estimated to be  $5.82 \cdot 10^{10}$  (see basal cell carcinoma section). We assume that each melanocyte is itself a stem cell because there is no evidence for a differentiation hierarchy in the melanocyte lineage. Melanocytes have been estimated to divide every  $\sim 147$  days (120), i.e. 2.48 times per year.

### Osteosarcoma

The lifetime incidence of bone and joint cancer is 0.1% (www.seer.cancer.gov) (3). Approximately 35% of bone and joint cancers are osteosarcomas (2). The proportion of osteosarcomas in the legs has been estimated to be 63.8% (18.5% tibia, 2.8% fibula, 42.5% femur), while 8.56% occur in the pelvis, 11.3% in the arms, and 8.63% in the head (skull and jaws) (2). As the lifetime incidence of osteosarcomas is  $0.001 \cdot 0.35 = 0.00035$ , it is  $0.001 \cdot 0.35 \cdot 0.638 = 0.00022$  for the legs,  $0.001 \cdot 0.35 \cdot 0.0856 = 0.00003$  for the pelvis,  $0.001 \cdot 0.35 \cdot 0.0113 = 0.00004$  for the arms, and  $0.001 \cdot 0.35 \cdot 0.0863 = 0.0000302$  for the head. In the average human body there are  $\sim 1.8 \cdot 10^9$  osteocytes (32). Osteocytes compose  $\sim 95\%$  of all bone cells in the adult skeleton, while osteoblasts comprise  $\sim 4\%$  and osteoclasts  $< 1\%$  (122). Thus there are  $\sim (4/95) \cdot 1.8 \cdot 10^9 = 7.6 \cdot 10^7$  osteoblasts and  $\sim 4-7\%$  of these have been estimated to be stem cells (123, 124). Thus we estimate that there are  $\sim 5.5\%$  of  $7.6 \cdot 10^7 = 4.18 \cdot 10^6$  osteocyte stem cells in the human skeleton. To calculate the numbers of osteocytes and osteocyte stem cells in different skeletal appendages, we used the data provided by USDHHS-NCHS 2013 for fractional bone content: 38% for the legs, 10.7% for the pelvis, 15.6% for the arms, and 20.65% for the head. Finally, bone stem cells divide every  $\sim 15$  years (81).

### Ovarian germ cell cancer

The lifetime incidence of ovarian cancer is  $\sim 1.37\%$  (8), and ovarian germ cell malignancies represent  $\sim 3\%$  of all ovarian cancers (see (125) and references therein). Thus, the lifetime incidence of ovarian germ cell cancer is  $0.03 \cdot 0.0137 = 0.000411$ . It has been shown in multiple studies that the number of germ cells (oocytes) in fetal ovaries reaches a peak of  $5.5 \cdot 10^6$  per gonad at 14-15 weeks of age p.c. to then declines, particularly after birth (126). The total number of germ cells present in both ovaries is then  $1.1 \cdot 10^7$ . There does not appear to be any oocyte divisions after week 15 of embryogenesis (127).

### Pancreatic ductal adenocarcinoma

The lifetime incidence of pancreatic cancer is 1.49% (www.seer.cancer.gov) (3). Exocrine cancers make up more than 96% of pancreatic cancers, with ductal adenocarcinoma being the most common type (95%) (www.cancer.org) (8). Thus, the lifetime incidence of pancreatic ductal adenocarcinoma is  $1.49\% \cdot 0.96 \cdot 0.95 = 1.3589\%$ . The pancreas contains Islets of Langerhans cells as well as acinar and ductal cells. Current evidence indicates that the cell of origin of pancreatic ductal adenocarcinomas is

the acinar cell (128), comprising 80-90% of the total pancreatic cells, while islet cells are  $2.95 \cdot 10^9$  in the pancreas (32), accounting for 1-2% of the total number of cells in the pancreas (129, 130). Thus, we estimate that the total number of acinar cells in the pancreas is  $((2.95 \cdot 10^9) / 1.5) \cdot 85 = 1.672 \cdot 10^{11}$ . Pancreatic stem cells (Bmi1-labeled cell lineage of pancreatic cells) maintaining organ homeostasis represent 2.5% of all pancreatic cells (130). Thus we estimate that the number of pancreatic stem cells is  $0.025 \cdot 1.672 \cdot 10^{11} = 4.18 \cdot 10^9$ . These stem cells are estimated to divide about once a year (131).

#### Pancreatic endocrine (islet cell) carcinoma

The lifetime incidence of pancreatic cancer is 1.49% (www.seer.cancer.gov) (3). Endocrine (islet cell) carcinomas account for about 1.3% of all pancreatic cancers (132). Thus, the lifetime incidence of pancreatic endocrine (islet cell) carcinoma is  $0.0149 \cdot 0.013 = 0.000194$ . The mean number of islet cells in a pancreas is equal to  $2.95 \cdot 10^9$ , representing 1 to 2% of the total pancreatic cells (32). Current evidence suggests that acinar and islet cells originate from a common progenitor (130, 133, 134). As noted above, pancreatic stem cells (Bmi1-labeled cell lineage of pancreatic cells) maintaining organ homeostasis represent 2.5% of all pancreatic cells (130). Thus we estimate that the number of islet stem cells is  $0.025 \cdot 2.95 \cdot 10^9 = 7.4 \cdot 10^7$ . These stem cells are estimated to divide about once a year (131, 135).

#### Small intestinal adenocarcinoma

The lifetime incidence of small intestine cancer is ~0.2% (www.seer.cancer.gov) (3), and ~35% of these cancers are adenocarcinomas (8). Thus, the lifetime incidence of small intestine adenocarcinoma is  $\sim 0.002 \cdot 0.35 = 0.0007$ . The small intestine is on average 6 meters long and ~2.5cm in diameter, for a total surface of approximately  $4,700 \text{ cm}^2$  (136), that is, about 1.4 times the surface of the large intestine. The total number of cells in the small intestine is  $\sim 1.7 \cdot 10^{10}$  (32). While the small intestine has crypts as in the large intestine, the crypts of the large intestine are more abundant, deeper, and have higher rates cell turnover than those of the small intestine. For example, it has been estimated that the crypts are ~0.3 mm deep in the duodenum (136, 137), ~0.16mm in the jejunum and ~0.175mm in the ileum (138), while they are 0.4-0.6mm deep in the large intestine (136). More importantly, the density of the crypts is ~100 crypts/mm<sup>2</sup> in the jejunum and ~50 crypts/mm<sup>2</sup> in the ileum (138), while it is ~140 crypts/mm<sup>2</sup> in the large intestine (52, 58). Note that the villi in the small intestine occupy a large amount of surface area, explaining the lower crypt density. Also, the stem cells per crypt are fewer in the small intestine; the stem-cells-per-crypt ratio between the small intestine and the large intestine is 5:7 for the proximal small intestine and 6:7 in the distal small intestine (65). The duodenum, jejunum and ileum are 0.25 m, 2.5, and 3.5 m, respectively, and the total number of crypts can therefore be calculated to be  $1.4 \cdot 1.5 \cdot 10^7 \cdot ((100 \cdot 2.5/6 + 50 \cdot 3.5/6)/140) \cdot (5.5/7) \sim 8 \cdot 10^6$ . Similarly, the total number of stem cells in the small intestine can be calculated to be approximately  $1.4 \cdot 2 \cdot 10^8 \cdot ((100 \cdot 2.5/6 + 50 \cdot 3.5/6)/140) \cdot (5.5/7) \sim 1 \cdot 10^8$ , which is very close to what was previously estimated by Potten *et al.* in (55). Finally, stem cells in the duodenum and distal small intestine divide, respectively, only 1/3 and 2/3 as frequently as those in the large intestine (64, 65). Thus, in humans, the small intestine undergoes fewer stem cell divisions than the large intestine (see Table

1). Interestingly, in mice the opposite is true: mouse small intestine undergoes more stem cell divisions than mouse large intestine (55, 139).

### Testicular germ cell cancer

The lifetime incidence of testicular cancer is ~0.39% (8), and ~95% of these cancers originate from germ cells (140). Thus, the lifetime incidence of testicular germ cell cancer is  $\sim 0.0039 \cdot 0.95 = 0.0037$ . The number of spermatogonial stem cells ( $A_s$ , spermatogonia) has been estimated to be ~35,000 in a mouse testis, and 350,000 in a rat testis (with the rat testis weighting ~10 times more than a mouse testis) (141-143). As the average human testis weighs ~17 grams (144, 145), comparison with the rat indicates that the average number of human spermatogonial stem cells (type  $A_d$ ) in a human is  $2 \cdot 350,000 \cdot 17/1.65 = 7.2 \cdot 10^6$  (see also (126)). As there are ~2,000 to 4,000 spermatozoa for every spermatogonial stem cell (stem cell fraction of ~0.03%) (142, 143), the average number of spermatozoa present in a human male can be calculated to be  $\sim 2.16 \cdot 10^{10}$ . Because  $\sim 10^{13}$  spermatozoa are produced in a lifetime (146), there is an average of  $10^{13} / (2.16 \cdot 10^{10}) = 463$  complete turnovers in a lifetime, implying that spermatogonial stem cells divide  $\sim 463/80 = 5.8$  times per year. This is consistent with the observation that the whole process of human spermatogenesis requires ~2-3 months (145).

### Thyroid papillary and follicular carcinoma

The lifetime incidence of thyroid cancer is ~1.08% (www.seer.cancer.gov) (3), and more than 95% of these cancers (both papillary and follicular) are derived from follicular cells (106-108). Thus, the combined lifetime incidence of papillary and follicular thyroid carcinomas is  $\sim 0.0108 \cdot 0.95 = 0.01026$ . There are  $\sim 10^{10}$  follicular cells in the thyroid. The stem cell population has been estimated to represent ~0.1-1.4% of the total cells (112, 114-116). Thus we assume there are  $0.0065 \cdot 10^{10} = 6.5 \cdot 10^7$  stem cells in the thyroid. Cell turnover has been estimated to be very low, with stem cell divisions occurring once every 8.5 – 14.4 years (116, 117).

### Statistical Analysis

The total number of stem cell divisions in the lifetime of a tissue was calculated as follows. Let  $s$  be the total number of stem cells found in a fully developed tissue, with  $s$  a power of 2, for simplicity, and no cell death. Starting from the first precursor cell of that tissue, it takes  $x$  generations during development to generate all of these cells, where  $2^x = s$ . Once the tissue has been fully developed, each of these  $s$  cells undergoes a total of  $d$  further divisions, due to normal tissue turnover, in the lifetime of that tissue. These turnover divisions are assumed to be asymmetric, but note that a balance between apoptosis and symmetric self-renewal would yield the same average number of cell divisions for a tissue in homeostasis. Thus, the cumulative number of division events, each yielding a new stem cell, among all stem cells in a lifetime ( $lscd$ , for lifetime stem cells divisions), is

$$lscd = \sum_{n=1}^{\log_2 s} 2^n + s \cdot d.$$

In general,  $s$  is not a power of 2, and the use of the floor function to approximate  $\log_2 s$  may not be appropriate. Noting that the partial sum of the geometric series is equal to  $2s-2$ , we obtain our formula for  $lscd$  for a general  $s$ :

$$lscd = s(2 + d) - 2.$$

The estimates for  $s$  and  $d$  are provided in Table 1. For each cancer type,  $lscd = s(2 + d) - 2$  was plotted against the lifetime incidence of that specific cancer type in Fig. 1 of the main text.

All statistical analyses were performed using *R*, version 3.1.0. Spearman's correlation test yielded a highly significant ( $P < 3.5 \cdot 10^{-8}$ ) correlation 0.81 (0.61 – 0.91; 95% CI), where the 95% confidence interval was calculated by bootstrap. Importantly, for the same data points in the log-log plot of Fig. 1, Pearson's yielded equivalent results with a highly significant linear correlation ( $P < 5.15 \cdot 10^{-8}$ ), and with a correlation coefficient equal to 0.804 (0.63-0.90; 95% CI). The coefficient of determination (R-squared) is therefore equal to 0.646 (0.395 – 0.813; 95% CI). The rationale behind calculating a p-value is as follows. The sample space is the set of all human tissue types, where each tissue has a lifetime cancer incidence value and a  $lscd$  value. Importantly, this sample space is large. The tissue types for which the needed estimates exist are a random (and small) subset of the sample space (their availability depends on other researchers' work).

The robustness of the correlation illustrated in Fig. 1 was tested by allowing the estimates for the number of stem cell divisions plotted in Fig. 1 (derived from the estimates in Table 1), to vary by ~100-fold in either direction. Specifically, first we sampled from a normal distribution with standard deviation 1 and mean equal to  $\log_{10}(lscd)$  for each x-coordinate value of the 31 data points in Fig. 1. Then Spearman's was used to test the correlation. After 10,000 iteration of this process, the median positive correlation was 0.7 (total range: 0.41 – 0.89), with a median p-value  $< 2 \cdot 10^{-5}$ . Importantly, all iterations yielded a statistically significant positive correlation, with the highest p-value  $< 2.4 \cdot 10^{-2}$ . Even replacing the normal distribution with the more extreme uniform distribution over the interval  $(-2 + \log_{10}(lscd), 2 + \log_{10}(lscd))$  for each data point yielded a significant positive correlation over 10,000 iterations with a median correlation 0.67 (total range: 0.39 – 0.90) (highest p-value  $< 2.9 \cdot 10^{-2}$ ). Thus, though the total range for  $lscd$  is ~ 6 orders of magnitude and we allowed four 4 orders of magnitude variation for each data point, the correlations generated were always statistically significant. This provides strong evidence that our results are robust.

For each cancer type (data points of Fig. 1), we defined an Extra Risk Score (ERS) to be the product of the  $\log_{10}$  value of lifetime cancer incidence,  $r$ , and its lifetime number of stem cell divisions,  $lscd$  (the values for  $r$  and  $lscd$  are provided in Table 1):

$$ERS = \log_{10} r \cdot \log_{10} lscd .$$

Note that  $\log_{10} lscd$  and  $\log_{10} r$  are simply the  $\log_{10}$  of the x- and y-coordinates of each point in Fig 1, and that the greater the absolute value of this product is, the smaller the

evidence for the presence of any environmental or inherited factor acting on that tissue. ERS represents the (negative value of the) area of the rectangle formed in the upper-left quadrant of Fig. 1 by the two coordinates (in logarithmic scale) of a data point as its sides. The larger the area of this rectangle, i.e. the more negative ERS, the less evidence for external environmental or inherited factors adding their extra effects to the stochastic replicative ones. Note that using the ratio between the  $\log_{10}$  values of  $r$  and  $lscd$ , instead of the product, would be sub-optimal to estimate the extra risk. For example, a cancer type with a lifetime incidence of  $10^{-6}$  and a total number of stem cell divisions of  $10^{12}$  would have the same *ratio* as a cancer with a lifetime incidence of  $10^{-3}$  and a total number of stem cell divisions of  $10^6$  (-6/12 for the first cancer and -3/6 for the second; both = -0.5). But any score reflecting the extra risk (i.e., the risk over and above that associated with cell divisions) should be much lower for the first cancer type than the second, because the first cancer type has a lower incidence despite a much higher number of stem cell divisions. When ERS is defined as the product rather than the ratio, the expected relationship is evident: the ERS for the first cancer type is -72, while the ERS for the second cancer type is much higher (i.e., -18).

An unsupervised machine learning analysis was performed and K-means cluster analysis was applied to the ERS scores, yielding the clustering depicted in Fig. 2. As it is well known that there is no established way to ascertain the “correct” number of clusters, and given that our primary interest was in identifying the cancer types in which strong environmental or inherited factors played a role, the number of clusters was set to two. Increasing the number of clusters yielded the expected result. For example, if the number of clusters was set to three, the cluster defined as “D-tumors” split in two, one of which contains cancer types where environmental or inherited factors are known to play a role, and another where these effects are not known to play a role. Thus, cancer types may be viewed as forming a continuum, from the most affected by environmental and inherited factors to the least, and the ERS score may be used to rank them accordingly. Ward hierarchical clustering, performed via the *hclust* function in the *stats R* package, yielded comparable results, with the three R-tumors closest to the D-tumors shifting to the D-tumor cluster (compare Fig. 2 with Fig. S1).

Given the results of this clustering, where the smallest ERS value in the D-tumors cluster was approximately -17.44 and the largest value in the R-tumors cluster was -19.54, we defined the “adjusted Extra Risk Score” (aERS) as  $aERS = 18.49 + ERS$ , where 18.49 is the average between the absolute values of the above two numbers. The aERS values are provided in Figure 2 of the main text.

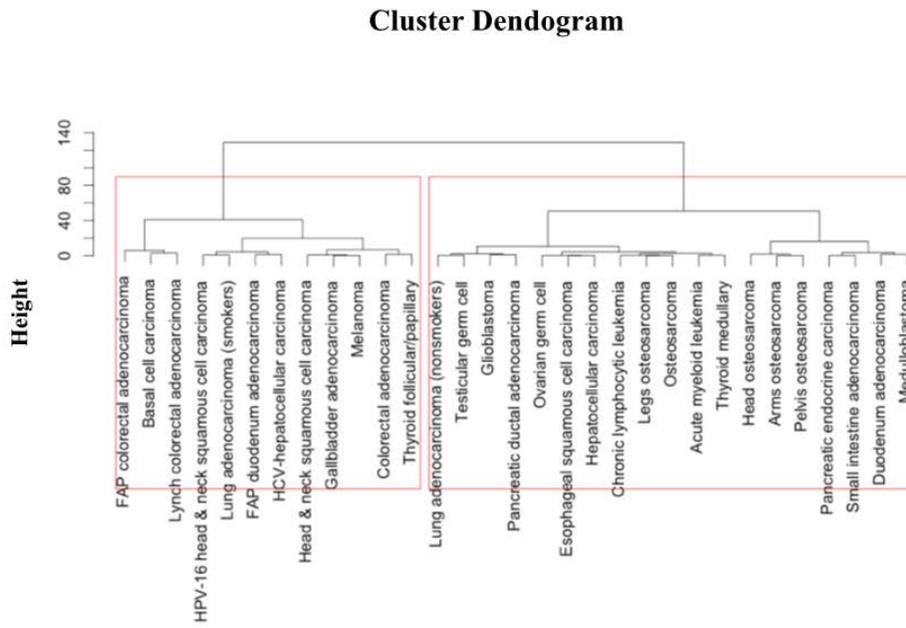


Fig. S1. Ward hierarchical clustering of D- and R-tumors.

**Table S1.**

Lifetime cancer incidence and parameters related to the normal stem cells that are precursors of these cancers. “Stem cells” denote the self-renewing cell population responsible for the homeostasis of the relevant cells in the indicated tissues. Lifetime parameters were obtained from estimates of the frequency of tissue self-renewal, assuming an average lifespan of 80 years. The definitions for the parameters  $s$ ,  $d$ , and  $lscd$  are provided in the Statistical Analysis section of the Materials and Methods section. The literature sources for the parameters listed in this Table are provided for each cancer type in the Materials and Methods section.

Cancer type	Lifetime cancer incidence	Total number of normal cells* in tissue of origin	Number of normal stem cells* in tissue of origin (s)	Number of divisions of each stem cell per year	Number of divisions of each stem cell per lifetime (d)	Cumulative number of divisions of all stem cells per lifetime (lscd)
Acute myeloid leukemia	0.0041	$3 \cdot 10^{12}$	$1.35 \cdot 10^8$	12	960	$1.299 \cdot 10^{11}$
Basal cell carcinoma	0.3	$1.8 \cdot 10^{11}$	$5.82 \cdot 10^9$	7.6	608	$3.550 \cdot 10^{12}$
Chronic lymphocytic leukemia	0.0052	$3 \cdot 10^{12}$	$1.35 \cdot 10^8$	12	960	$1.299 \cdot 10^{11}$
Colorectal adenocarcinoma	0.048	$3 \cdot 10^{10}$	$2 \cdot 10^8$	73	5840	$1.168 \cdot 10^{12}$
Colorectal adenocarcinoma with FAP	1	$3 \cdot 10^{10}$	$2 \cdot 10^8$	73	5840	$1.168 \cdot 10^{12}$
Colorectal adenocarcinoma with Lynch syndrome	0.5	$3 \cdot 10^{10}$	$2 \cdot 10^8$	73	5840	$1.168 \cdot 10^{12}$
Duodenum adenocarcinoma	0.0003	$6.8 \cdot 10^8$	$4 \cdot 10^6$	24	1947	$7.796 \cdot 10^9$

Duodenum adenocarcinoma with FAP	0.035	$6.8 \cdot 10^8$	$4 \cdot 10^6$	24	1947	$7.796 \cdot 10^9$
Esophageal squamous cell carcinoma	0.001938	$3.24 \cdot 10^9$	$8.64 \cdot 10^5$	17.4	1390	$1.203 \cdot 10^9$
Gallbladder non papillary adenocarcinoma	0.0028	$1.6 \cdot 10^8$	$1.6 \cdot 10^6$	0.584	47	$7.840 \cdot 10^7$
Glioblastoma	0.00219	$8.46 \cdot 10^{10}$	$1.35 \cdot 10^8$	0	0	$2.700 \cdot 10^8$
Head & neck squamous cell carcinoma	0.0138	$1.67 \cdot 10^{10}$	$1.85 \cdot 10^7$	21.5	1720	$3.186 \cdot 10^{10}$
Head & neck squamous cell carcinoma with HPV-16	0.07935	$1.67 \cdot 10^{10}$	$1.85 \cdot 10^7$	21.5	1720	$3.186 \cdot 10^{10}$
Hepatocellular carcinoma	0.0071	$2.41 \cdot 10^{11}$	$3.01 \cdot 10^9$	0.9125	88	$2.709 \cdot 10^{11}$
Hepatocellular carcinoma with HCV	0.071	$2.41 \cdot 10^{11}$	$3.01 \cdot 10^9$	0.9125	88	$2.709 \cdot 10^{11}$
Lung adenocarcinoma (nonsmokers)	0.0045	$4.34 \cdot 10^{11}$	$1.22 \cdot 10^9$	0.07	5.6	$9.272 \cdot 10^9$
Lung adenocarcinoma (smokers)	0.081	$4.34 \cdot 10^{11}$	$1.22 \cdot 10^9$	0.07	5.6	$9.272 \cdot 10^9$
Medulloblastoma	0.00011	$8.5 \cdot 10^{10}$	$1.36 \cdot 10^8$	0	0	$2.720 \cdot 10^8$
Melanoma	0.0203	$3.8 \cdot 10^9$	$3.8 \cdot 10^9$	2.48	199	$7.638 \cdot 10^{11}$
Osteosarcoma	0.00035	$1.9 \cdot 10^9$	$4.18 \cdot 10^6$	0.067	5	$2.926 \cdot 10^7$
Osteosarcoma of the arms	0.00004	$3 \cdot 10^8$	$6.5 \cdot 10^5$	0.067	5	$4.550 \cdot 10^6$
Osteosarcoma of the head	0.000030 2	$3.9 \cdot 10^8$	$8.6 \cdot 10^5$	0.067	5	$6.020 \cdot 10^6$
Osteosarcoma of	0.00022	$7.2 \cdot 10^8$	$1.59 \cdot 10^6$	0.067	5	$1.113 \cdot 10^7$

the legs						
Osteosarcoma of the pelvis	0.00003	$2 \cdot 10^8$	$4.5 \cdot 10^5$	0.067	5	$3.150 \cdot 10^6$
Ovarian germ cell	0.000411	$1.1 \cdot 10^7$	$1.1 \cdot 10^7$	0	0	$2.200 \cdot 10^7$
Pancreatic ductal adenocarcinoma	0.013589	$1.672 \cdot 10^{11}$ (acinar)	$4.18 \cdot 10^9$	1	80	$3.428 \cdot 10^{11}$
Pancreatic endocrine (islet cell) carcinoma	0.000194	$2.95 \cdot 10^9$ (islet)	$7.4 \cdot 10^7$	1	80	$6.068 \cdot 10^9$
Small intestine adenocarcinoma	0.0007	$1.7 \cdot 10^{10}$	$1 \cdot 10^8$	36	2920	$2.922 \cdot 10^{11}$
Testicular germ cell cancer	0.0037	$2.16 \cdot 10^{10}$	$7.2 \cdot 10^6$	5.8	463	$3.348 \cdot 10^9$
Thyroid papillary/follicular carcinoma	0.01026	$10^{10}$	$6.5 \cdot 10^7$	0.087	7	$5.850 \cdot 10^8$
Thyroid medullary carcinoma	0.000324	$10^9$	$6.5 \cdot 10^6$	0.087	7	$5.850 \cdot 10^7$

\*"Cells" and "stem cells" refer only to those normal cells of the same type as the cancer cells in that tissue. For example, for colorectal adenocarcinomas, the cells and stem cells referred to are epithelial cells, not the stromal or other cell types within normal colon. For some cancer types, such as osteosarcomas, overall data as well as anatomic-compartment specific data are included.

## References and Notes

32. E. Bianconi *et al.*, An estimation of the number of cells in the human body. *Ann Hum Biol* **40**, 463 (Nov-Dec, 2013).
33. K. Kato, A. Radbruch, Isolation and characterization of CD34+ hematopoietic stem cells from human peripheral blood by high-gradient magnetic cell sorting. *Cytometry* **14**, 384 (1993).
34. W. W. Pang *et al.*, Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc Natl Acad Sci U S A* **108**, 20012 (Dec 13, 2011).
35. A. Foudi *et al.*, Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells. *Nat Biotechnol* **27**, 84 (Jan, 2009).
36. M. J. Kiel *et al.*, Haematopoietic stem cells do not asymmetrically segregate chromosomes or retain BrdU. *Nature* **449**, 238 (Sep 13, 2007).
37. G. B. Bradford, B. Williams, R. Rossi, I. Bertoncello, Quiescence, cycling, and turnover in the primitive hematopoietic stem cell compartment. *Exp Hematol* **25**, 445 (May, 1997).
38. K. Sudo, H. Ema, Y. Morita, H. Nakauchi, Age-associated characteristics of murine hematopoietic stem cells. *J Exp Med* **192**, 1273 (Nov 6, 2000).
39. S. H. Cheshier, S. J. Morrison, X. Liao, I. L. Weissman, In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells. *Proc Natl Acad Sci U S A* **96**, 3120 (Mar 16, 1999).
40. T. L. Diepgen, V. Mahler, The epidemiology of skin cancer. *Br J Dermatol* **146 Suppl 61**, 1 (Apr, 2002).
41. D. L. Miller, M. A. Weinstock, Nonmelanoma skin cancer in the United States: incidence. *J Am Acad Dermatol* **30**, 774 (May, 1994).
42. A. I. Rubin, E. H. Chen, D. Ratner, Basal-cell carcinoma. *N Engl J Med* **353**, 2262 (Nov 24, 2005).
43. M. Kasper, V. Jaks, D. Hohl, R. Toftgard, Basal cell carcinoma - molecular biology and potential new therapies. *J Clin Invest* **122**, 455 (Feb 1, 2012).
44. K. K. Youssef *et al.*, Identification of the cell lineage at the origin of basal cell carcinoma. *Nat Cell Biol* **12**, 299 (Mar, 2010).
45. A. Rook, T. Burns, *Rook's textbook of dermatology*. (Blackwell Science, Malden, Mass., ed. 7th, 2004).
46. J. Nyman, I. Turesson, Basal cell density in human skin for various fractionation schedules in radiotherapy. *Radiother Oncol* **33**, 117 (Nov, 1994).
47. U. B. Jensen, S. Lowell, F. M. Watt, The spatial relationship between stem cells and their progeny in the basal layer of human epidermis: a new view based on whole-mount labelling and lineage analysis. *Development* **126**, 2409 (Jun, 1999).
48. P. H. Jones, S. Harper, F. M. Watt, Stem cell patterning and fate in human epidermis. *Cell* **80**, 83 (Jan 13, 1995).
49. R. D. Mosteller, Simplified calculation of body-surface area. *N Engl J Med* **317**, 1098 (Oct 22, 1987).

50. K. M. Halprin, Epidermal "turnover time"--a re-examination. *Br J Dermatol* **86**, 14 (Jan, 1972).
51. H. Iizuka, Epidermal turnover time. *J Dermatol Sci* **8**, 215 (Dec, 1994).
52. B. M. Boman, J. Z. Fields, K. L. Cavanaugh, A. Guetter, O. A. Runquist, How dysregulated colonic crypt dynamics cause stem cell overpopulation and initiate colon cancer. *Cancer Res* **68**, 3304 (May 1, 2008).
53. P. Calabrese, S. Tavaré, D. Shibata, Pretumor progression: clonal evolution of human stem cell populations. *Am J Pathol* **164**, 1337 (Apr, 2004).
54. A. Facista *et al.*, Deficient expression of DNA repair enzymes in early progression to sporadic colon cancer. *Genome Integr* **3**, 3 (2012).
55. C. S. Potten, C. Booth, D. Hargreaves, The small intestine as a model for evaluating adult tissue stem cell drug targets. *Cell Prolif* **36**, 115 (Jun, 2003).
56. J. H. Song *et al.*, The APC network regulates the removal of mutated cells from colonic crypts. *Cell Rep* **7**, 94 (Apr 10, 2014).
57. P. Nicolas, K. M. Kim, D. Shibata, S. Tavaré, The stem cell population of the human colon crypt: analysis via methylation patterns. *PLoS Comput Biol* **3**, e28 (Mar 2, 2007).
58. C. S. Potten, M. Kellett, S. A. Roberts, D. A. Rew, G. D. Wilson, Measurement of in vivo proliferation in human colorectal mucosa using bromodeoxyuridine. *Gut* **33**, 71 (Jan, 1992).
59. H. Cheng, M. Bjerknes, J. Amar, Methods for the determination of epithelial cell kinetic parameters of human colonic epithelium isolated from surgical and biopsy specimens. *Gastroenterology* **86**, 78 (Jan, 1984).
60. R. Okamoto, M. Watanabe, Molecular and clinical basis for the regeneration of human gastrointestinal epithelia. *J Gastroenterol* **39**, 1 (Jan, 2004).
61. H. T. Lynch, A. de la Chapelle, Hereditary colorectal cancer. *N Engl J Med* **348**, 919 (Mar 6, 2003).
62. J. G. Dowty *et al.*, Cancer risks for MLH1 and MSH2 mutation carriers. *Hum Mutat* **34**, 490 (Mar, 2013).
63. A. I. Neugut, J. S. Jacobson, S. Suh, R. Mukherjee, N. Arber, The epidemiology of cancer of the small bowel. *Cancer Epidemiol Biomarkers Prev* **7**, 243 (Mar, 1998).
64. J. Y. Kim, K. D. Siegmund, S. Tavaré, D. Shibata, Age-related human small intestine methylation: evidence for stem cell niches. *BMC Med* **3**, 10 (2005).
65. S. Kozar *et al.*, Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. *Cell Stem Cell* **13**, 626 (Nov 7, 2013).
66. K. F. Trivers, S. A. Sabatino, S. L. Stewart, Trends in esophageal cancer incidence by histology, United States, 1998-2003. *Int J Cancer* **123**, 1422 (Sep 15, 2008).
67. R. Kim, J. L. Weissfeld, J. C. Reynolds, L. H. Kuller, Etiology of Barrett's metaplasia and esophageal adenocarcinoma. *Cancer Epidemiol Biomarkers Prev* **6**, 369 (May, 1997).
68. M. H. Sleisenger, M. Feldman, L. S. Friedman, L. J. Brandt, *Sleisenger and Fordtran's gastrointestinal and liver disease : pathophysiology, diagnosis, management*. (Saunders/Elsevier, Philadelphia , PA, ed. 9th, 2010).

69. T. Smitha, P. Sharada, H. Girish, Morphometry of the basal cell layer of oral leukoplakia and oral squamous cell carcinoma using computer-aided image analysis. *J Oral Maxillofac Pathol* **15**, 26 (Jan, 2011).
70. D. P. Doupe *et al.*, A single progenitor population switches behavior to maintain and repair esophageal epithelium. *Science* **337**, 1091 (Aug 31, 2012).
71. C. A. Squier, K. A. Brogden, *Human oral mucosa : development, structure, and function*. (Wiley-Blackwell, Chichester, West Sussex, UK, 2011), pp. viii, 168 p.
72. C. A. Squier, M. J. Kremer, Biology of oral mucosa and esophagus. *J Natl Cancer Inst Monogr*, 7 (2001).
73. <http://www.cancerresearchuk.org>.
74. S. P. Lee, C. E. Savard, R. Kuver, Gallbladder epithelial cells that engraft in mouse liver can differentiate into hepatocyte-like cells. *Am J Pathol* **174**, 842 (Mar, 2009).
75. P. Putz, G. Willems, Cell proliferation in the human gallbladder epithelium: effect of distension. *Gut* **20**, 246 (Mar, 1979).
76. Q. T. Ostrom *et al.*, The epidemiology of glioma in adults: a "state of the science" review. *Neuro Oncol*, (May 19, 2014).
77. F. A. Azevedo *et al.*, Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J Comp Neurol* **513**, 532 (Apr 10, 2009).
78. S. Herculano-Houzel, The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proc Natl Acad Sci U S A* **109 Suppl 1**, 10661 (Jun 26, 2012).
79. C. Eroglu, B. A. Barres, Regulation of synaptic connectivity by glia. *Nature* **468**, 223 (Nov 11, 2010).
80. B. A. Reynolds, R. L. Rietze, Neural stem cells and neurospheres--re-evaluating the relationship. *Nat Methods* **2**, 333 (May, 2005).
81. K. L. Spalding, R. D. Bhardwaj, B. A. Buchholz, H. Druid, J. Frisen, Retrospective birth dating of cells in humans. *Cell* **122**, 133 (Jul 15, 2005).
82. G. D'Souza *et al.*, Case-control study of human papillomavirus and oropharyngeal cancer. *N Engl J Med* **356**, 1944 (May 10, 2007).
83. M. L. Gillison *et al.*, Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *J Natl Cancer Inst* **92**, 709 (May 3, 2000).
84. A. K. Chaturvedi *et al.*, Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **31**, 4550 (Dec 20, 2013).
85. J. Mork *et al.*, Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck. *N Engl J Med* **344**, 1125 (Apr 12, 2001).
86. X. Li *et al.*, Human papillomavirus infection and laryngeal cancer risk: a systematic review and meta-analysis. *J Infect Dis* **207**, 479 (Feb 1, 2013).
87. M. L. Gillison *et al.*, Prevalence of oral HPV infection in the United States, 2009-2010. *JAMA* **307**, 693 (Feb 15, 2012).
88. K. R. Dahlstrom *et al.*, Human papillomavirus type 16 infection and squamous cell carcinoma of the head and neck in never-smokers: a matched pair analysis. *Clin Cancer Res* **9**, 2620 (Jul, 2003).

89. E. A. Naumova, T. Dierkes, J. Sprang, W. H. Arnold, The oral mucosal surface and blood vessels. *Head Face Med* **9**, 8 (2013).
90. L. M. Collins, C. Dawes, The surface area of the adult human mouth and thickness of the salivary film covering the teeth and oral mucosa. *J Dent Res* **66**, 1300 (Aug, 1987).
91. M. M. Daniel, M. C. Lorenzi, C. da Costa Leite, G. Lorenzi-Filho, Pharyngeal dimensions in healthy men and women. *Clinics (Sao Paulo)* **62**, 5 (Feb, 2007).
92. H. E. Eckel, C. Sittel, P. Zorowka, A. Jerke, Dimensions of the laryngeal framework in adults. *Surg Radiol Anat* **16**, 31 (1994).
93. K. B. Jones, O. D. Klein, Oral epithelial stem cells in tissue maintenance and disease: the first steps in a long journey. *Int J Oral Sci* **5**, 121 (Sep, 2013).
94. A. H. Shojaei, Buccal mucosa as a route for systemic drug delivery: a review. *J Pharm Pharm Sci* **1**, 15 (Jan-Apr, 1998).
95. S. F. Altekruse, K. A. McGlynn, M. E. Reichman, Hepatocellular carcinoma incidence, mortality, and survival trends in the United States from 1975 to 2005. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **27**, 1485 (Mar 20, 2009).
96. J. A. Davila, R. O. Morgan, Y. Shaib, K. A. McGlynn, H. B. El-Serag, Hepatitis C infection and the increasing incidence of hepatocellular carcinoma: a population-based study. *Gastroenterology* **127**, 1372 (Nov, 2004).
97. G. L. Armstrong *et al.*, The prevalence of hepatitis C virus infection in the United States, 1999 through 2002. *Ann Intern Med* **144**, 705 (May 16, 2006).
98. R. Turner *et al.*, Human hepatic stem cell and maturational liver lineage biology. *Hepatology* **53**, 1035 (Mar, 2011).
99. W. K. Hong, American Association for Cancer Research., *Holland Frei cancer medicine* 8. (People's Medical Pub. House, Shelton, Conn., ed. 8th, 2010), pp. xxv, 2,021 p.
100. "Current cigarette smoking among adults - United States, 2005 - 2012. " *Morbidity and Mortality Weekly Report*. 62(02):29-34 (Centers for Disease Control and Prevention, 2014).
101. C. F. Kim *et al.*, Identification of bronchioalveolar stem cells in normal lung and lung cancer. *Cell* **121**, 823 (Jun 17, 2005).
102. J. H. Lee *et al.*, Lung stem cell differentiation in mice directed by endothelial cells via a BMP4-NFATc1-thrombospondin-1 axis. *Cell* **156**, 440 (Jan 30, 2014).
103. T. J. Desai, D. G. Brownfield, M. A. Krasnow, Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* **507**, 190 (Mar 13, 2014).
104. T. A. Dolecek, J. M. Propp, N. E. Stroup, C. Kruchko, CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2005-2009. *Neuro Oncol* **14 Suppl 5**, v1 (Nov, 2012).
105. R. J. Gilbertson, D. W. Ellison, The origins of medulloblastoma subtypes. *Annu Rev Pathol* **3**, 341 (2008).
106. T. Kondo, S. Ezzat, S. L. Asa, Pathogenetic mechanisms in thyroid follicular-cell neoplasia. *Nat Rev Cancer* **6**, 292 (Apr, 2006).
107. S. A. Hundahl, I. D. Fleming, A. M. Fremgen, H. R. Menck, A National Cancer Data Base report on 53,856 cases of thyroid carcinoma treated in the U.S., 1985-1995 [see commetns]. *Cancer* **83**, 2638 (Dec 15, 1998).

108. V. A. LiVolsi, S. L. Asa, The demise of follicular carcinoma of the thyroid gland. *Thyroid* **4**, 233 (Summer, 1994).
109. I. Martin-Lacave *et al.*, C cells evolve at the same rhythm as follicular cells when thyroidal status changes in rats. *J Anat* **214**, 301 (Mar, 2009).
110. I. Martin-Lacave, E. Conde, C. Montero, H. Galera-Davidson, Quantitative changes in the frequency and distribution of the C-cell population in the rat thyroid gland with age. *Cell Tissue Res* **270**, 73 (Oct, 1992).
111. G. Lania *et al.*, Early thyroid development requires a Tbx1-Fgf8 pathway. *Dev Biol* **328**, 109 (Apr 1, 2009).
112. N. Hoshi, T. Kusakabe, B. J. Taylor, S. Kimura, Side population cells in the mouse thyroid exhibit stem/progenitor cell-like characteristics. *Endocrinology* **148**, 4251 (Sep, 2007).
113. T. Ozaki *et al.*, Thyroid regeneration: characterization of clear cells after partial thyroidectomy. *Endocrinology* **153**, 2514 (May, 2012).
114. R. Y. Lin, New insights into thyroid stem cells. *Thyroid* **17**, 1019 (Oct, 2007).
115. R. Y. Lin, Thyroid cancer stem cells. *Nat Rev Endocrinol* **7**, 609 (Oct, 2011).
116. A. Fierabracci, Identifying thyroid stem/progenitor cells: advances and limitations. *J Endocrinol* **213**, 1 (Apr, 2012).
117. J. Coclet, F. Foureau, P. Ketelbant, P. Galand, J. E. Dumont, Cell population kinetics in dog and human adult thyroid. *Clin Endocrinol (Oxf)* **31**, 655 (Dec, 1989).
118. A. J. Miller, M. C. Mihm, Jr., Melanoma. *N Engl J Med* **355**, 51 (Jul 6, 2006).
119. J. Kanitakis, Anatomy, histology and immunohistochemistry of normal human skin. *Eur J Dermatol* **12**, 390 (Jul-Aug, 2002).
120. K. Jimbow, S. I. Roth, T. B. Fitzpatrick, G. Szabo, Mitotic activity in non-neoplastic melanocytes in vivo as determined by histochemical, autoradiographic, and electron microscope studies. *J Cell Biol* **66**, 663 (Sep, 1975).
121. H. Chung, E. K. Suh, I. O. Han, E. S. Oh, Keratinocyte-derived laminin-332 promotes adhesion and migration in melanocytes and melanoma. *J Biol Chem* **286**, 13438 (Apr 15, 2011).
122. L. F. Bonewald, Osteocytes as dynamic multifunctional cells. *Ann N Y Acad Sci* **1116**, 281 (Nov, 2007).
123. R. W. Young, Cell proliferation and specialization during endochondral osteogenesis in young rats. *J Cell Biol* **14**, 357 (Sep, 1962).
124. A. Cicconetti *et al.*, Human maxillary tuberosity and jaw periosteum as sources of osteoprogenitor cells for tissue engineering. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* **104**, 618 e1 (Nov, 2007).
125. R. S. Arora, R. D. Alston, T. O. B. Eden, M. Geraci, J. M. Birch, Comparative incidence patterns and trends of gonadal and extragonadal germ cell tumors in England, 1979 to 2003. *Cancer* **118**, 4290 (Sep 1, 2012).
126. L. S. Mamsen, M. C. Lutterodt, E. W. Andersen, A. G. Byskov, C. Y. Andersen, Germ cell numbers in human embryonic and fetal gonads during the first two trimesters of pregnancy: analysis of six published studies. *Human Reproduction* **26**, 2140 (Aug, 2011).
127. J. B. Kerr *et al.*, The primordial follicle reserve is not renewed after chemical or gamma-irradiation mediated depletion. *Reproduction* **143**, 469 (Apr, 2012).

128. J. L. Kopp *et al.*, Identification of Sox9-dependent acinar-to-ductal reprogramming as the principal mechanism for initiation of pancreatic ductal adenocarcinoma. *Cancer Cell* **22**, 737 (Dec 11, 2012).
129. R. Rubin, D. S. Strayer, E. Rubin, *Rubin's Pathology: Clinicopathologic Foundations of Medicine*. (Lippincott Williams & Wilkins, ed. 6th, 2011).
130. E. Sangiorgi, M. R. Capecchi, Bmi1 lineage tracing identifies a self-renewing pancreatic acinar cell subpopulation capable of maintaining pancreatic organ homeostasis. *Proc Natl Acad Sci U S A* **106**, 7101 (Apr 28, 2009).
131. K. Furuyama *et al.*, Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. *Nat Genet* **43**, 34 (Jan, 2011).
132. J. C. Yao *et al.*, Population-based study of islet cell carcinoma. *Ann Surg Oncol* **14**, 3492 (Dec, 2007).
133. A. Inada *et al.*, Carbonic anhydrase II-positive pancreatic cells are progenitors for both endocrine and exocrine pancreas after birth. *Proc Natl Acad Sci U S A* **105**, 19915 (Dec 16, 2008).
134. M. Rovira *et al.*, Isolation and characterization of centroacinar/terminal ductal progenitor cells in adult mouse pancreas. *Proc Natl Acad Sci U S A* **107**, 75 (Jan 5, 2010).
135. B. Tyrberg, J. Ustinov, T. Otonkoski, A. Andersson, Stimulated endocrine cell proliferation and differentiation in transplanted human pancreatic islets: effects of the ob gene and compensatory growth of the implantation organ. *Diabetes* **50**, 301 (Feb, 2001).
136. W. Kuehnel, *Pocket Atlas of Cytology, Histology and Microscopic Anatomy*. (Thieme, 2003).
137. A. G. Cummins *et al.*, Morphometric evaluation of duodenal biopsies in celiac disease. *Am J Gastroenterol* **106**, 145 (Jan, 2011).
138. J. B. Trbojevic-Stankovic *et al.*, Morphometric study of healthy jejunal and ileal mucosa in adult and aged subjects. *Histol Histopathol* **25**, 153 (Feb, 2010).
139. S. Behjati *et al.*, Genome sequencing of normal cells reveals developmental lineages and mutational processes. *Nature* **513**, 422 (Sep 18, 2014).
140. G. J. Bosl, R. J. Motzer, Testicular germ-cell cancer. *N Engl J Med* **337**, 242 (Jul 24, 1997).
141. S. Meachem, V. von Schonfeldt, S. Schlatt, Spermatogonia: stem cells with a great perspective. *Reproduction* **121**, 825 (Jun, 2001).
142. R. A. Tegelenbosch, D. G. de Rooij, A quantitative study of spermatogonial multiplication and stem cell renewal in the C3H/101 F1 hybrid mouse. *Mutat Res* **290**, 193 (Dec, 1993).
143. D. G. de Rooij, L. D. Russell, All you wanted to know about spermatogonia but were afraid to ask. *J Androl* **21**, 776 (Nov-Dec, 2000).
144. C. Mori *et al.*, Temporal changes in testis weight during the past 50 years in Japan. *Anat Sci Int* **77**, 109 (Jun, 2002).
145. R. P. Amann, Considerations in evaluating human spermatogenesis on the basis of total sperm per ejaculate. *J Androl* **30**, 626 (Nov-Dec, 2009).
146. R. Reijo *et al.*, Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. *Nat Genet* **10**, 383 (Aug, 1995).