Published in final edited form as:

Cancer Lett. 2014 October 1; 352(2): 137–144. doi:10.1016/j.canlet.2014.05.021.

Mobile elements and viral integrations prompt considerations for bacterial DNA integration as a novel carcinogen

Kelly M. Robinson¹ and Julie C. Dunning Hotopp^{1,2,3,*}

¹Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, MD, USA

²Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD, USA

³Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, MD, USA

Abstract

Insertional mutagenesis has been repeatedly demonstrated in cancer genomes and has a role in oncogenesis. Mobile genetic elements can induce cancer development by random insertion into cancer related genes or by inducing translocations. L1s are typically implicated in cancers of an epithelial cell origin, while *Alu* elements have been implicated in leukemia as well as epithelial cell cancers. Likewise, viral infections have a significant role in cancer development predominantly through integration into the human genome and mutating or deregulating cancer related genes. Human papilloma virus is the best-known example of viral integrations contributing to carcinogenesis. However, hepatitis B virus, Epstein-Barr virus, and Merkel cell polyomavirus also integrate into the human genome and disrupt cancer related genes. Thus far, the role of microbes in cancer has primarily been attributed to mutations induced through chronic inflammation or toxins, as is the case with *Helicobacter pylori* and enterotoxigenic *Bacteroides fragilis*. We hypothesize that like mobile elements and viral DNA, bacterial and parasitic DNA may also integrate into the human somatic genome and be oncogenic. Until recently it was believed that bacterial DNA could not integrate into the human genome, but new evidence demonstrates that bacterial insertional mutagenesis may occur in cancer cells. Although this work does not show causation between bacterial insertions and cancer, it prompts more research in this area. Promising new sequencing technologies may reduce the risk of artifactual chimeric sequences, thus diminishing some of the challenges of identifying novel insertions in the somatic human genome.

Keywords

Lateral gene transfer; mobile elements; viruses; bacteria; parasites

© 2014 Elsevier Ireland Ltd. All rights reserved.

*To whom correspondence should be addressed: jdhotopp@som.umaryland.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflict of Interest Statement
None

1. Introduction

In 2013, 1.6 million people were estimated to develop cancer in the US, while 49% of men and 38% of women will develop cancer in their lifetime. Genetic changes, environmental factors, and infectious agents can all cause increased cell proliferation leading to cancer through abnormal epigenetic alterations, point mutations, translocations, and other modifications. Infectious agents and mobile elements that integrate into the human genome, in whole or in part, are an example of events that inflict such DNA damage [122]. Here, we explore whether bacterial integrations may also occur in tumors.

2. Role of mobile elements in carcinogenesis

2.1 Mobile elements basics

Mobile elements are possibly the best-studied insertional mutagens of the human genome. While currently inactive in the human genome [92], DNA-based transposons account for ~45% of the human genome [40]. RNA-based retrotransposons can be divided into classes based on the presence/absence of long terminal repeats (LTR). The LTR-retrotransposons, such as endogenous retroviruses, are not currently active in humans [79]. However, the non-LTR retrotransposons actively jump throughout the human genome acting as mutagens [40, 54, 79] (Figure 1A). These non-LTR retrotransposons consist of long interspersed elements 1 (LINE-1s or L1s), *Alu* elements (a type of short interspersed element, or SINE), and SVAs (a combination of SINE-R, variable number of tandem repeats, and *Alu*-like sequences). L1s are the only non-LTR retrotransposons that encode enzymes for retrotransposition [54], and as such, they mobilize themselves as well as *Alu* elements and SVAs [54]. Of the >500,000 L1s in the human genome, only a few are highly active and responsible for the majority of new insertions [13]. In every generation, L1s and *Alu* elements move to new genomic locations through germline retrotransposition [54]. Many disease-causing transpositions have been identified thus far including ones in breast cancer [76, 128], Apert Syndrome [91], hemophilia [47, 85], and colon cancer [77, 123]. While L1s and *Alu* elements tend to be inactive in somatic tissues, their-reactivation could aid in tumorigenesis [54].

2.2 L1 insertions and cancer

L1s have been implicated in tumorigenesis of various epithelial cell cancers [14], including lung cancer [40, 46]. A genomic comparison of 20 lung cancer samples with normal matched samples revealed 9 L1 insertions in only the tumor samples [46]. Six of these samples had 1–3 new L1 insertions/tumor and also showed an increase in hypomethylated DNA when compared to the matched normal samples [46], meaning the DNA in these samples had decreased methylation. Since host cells reduce transposition through increased methylation, hypomethylated tumor cells are more permissive to L1 transposons [46].

Colorectal cancer samples also have an increased rate of L1 insertions, some of which have disrupted genes with known cancer driver functions [123], like the *APC* gene [77]. More tests are necessary to resolve the exact effect of L1 insertions since they may be passenger mutations or may be directly related to tumor formation [46, 123]. It has recently been reported that the methylation status of L1 insertions in rectal cancer predicts the clinical

outcome of the patient, with individuals with L1 hypomethylation having shorter survival times and higher incidences of tumor recurrence [8].

2.3 Alu elements and cancer

Alu elements are the most successful type of transposon with >1 million copies in the human genome [60]. *Alu* insertions have been associated with familial adenomatous polyposis, breast cancer, acute myeloid leukemia, and hereditary non-polyposis colorectal cancer syndrome [40].

2.3.1 Alu elements can disrupt DNA repair—*Alu* elements have been found to mutate tumor suppressor genes that aid in DNA repair, like *BRCA1* and *BRCA2*, which are important for homologous recombination in DNA double strand break repair. Recombination between two similar *Alu* elements led to the loss of a 3-kbp region in exon 17 of the *BRCA1* gene and subsequent inactivation of *BRCA1* causing breast cancer [82]. Other breast cancer studies have demonstrated multiple *Alu*-mediated mutations in the *BRCA1* and *BRCA2* genes [104]. Likewise, hereditary non-polyposis colorectal cancer syndrome can occur when mutations arise in genes involved in the mismatch repair system, such as *MLH1* or *MLH2* [80]. Both of these genes have increased *Alu* concentrations in their introns with about 75% of the rearrangements in the *MLH2* gene caused by *Alu*-mediated non-allelic homologous recombination [65].

2.3.2 Alu elements and leukemia—In acute myeloid leukemia, *Alu*-mediated partial duplications frequently occur in the coding regions of the *MLL1* gene, most commonly resulting in duplicate exons 5 through 11 or 12 to fuse upstream of the original exon 5 [112]. *MLL1* is part of the mixed lineage leukemia gene family, which have histone methylase capabilities and are involved in *HOX* gene regulation [2]. Some experiments suggest that the wild-type allele can be suppressed, allowing for expression of the *MLL1* copy containing the *Alu*-mediated partial duplication. This in turn induces a leukemic phenotype [134] through altered protein structure and inactivation of inherent *MLL1* function [111]. Overall, transposable elements could play a major role in somatic and germline mutagenesis and more focused research in this area is necessary before rates of L1 and *Alu* element involvement in cancer development and progression can be estimated.

3. Viral integrations in cancer

3.1 Viruses involved in carcinogenesis

In 2002, viral infections were estimated to cause 12.1% of cancers [93]. Known viral carcinogens include EBV, hepatitis B virus (HBV), hepatitis C virus (HCV), human papilloma virus (HPV), Kaposi's sarcoma-associated herpesvirus (KSHV), and human T-cell lymphotropic virus type 1 [108]. Human immunodeficiency virus type-1 [108, 116], Merkel cell polyomavirus (MCV) [30, 108], and BK polyomavirus [1, 74] are also proposed to be associated with cancer. While both DNA and RNA viruses have been linked to carcinogenesis, only the DNA viruses discussed in this review are known to integrate into the host genome and contribute to carcinogenesis through insertional mutagenesis.

3.2 Human papilloma virus

The insertion of HPV DNA into human chromosomes is the single most important event leading to tumorigenesis in cervical cancer (Figure 1B). Clonal integration of HPV-16 or HPV-18, two of the “high-risk” strains, [130], occurs in 80–100% of cervical carcinoma tumors [18, 19, 75]. Normally, the HPV E2 protein regulates E6 and E7 oncogene expression [6] and while the virus replicates, the host cell maintains control of proliferation [133]. However, when HPV integrates into the host nuclear genome it frequently does so in a manner that leads to disruption of E2 [133]. The integration of E6 and E7 without a functional E2 deregulates E6 and E7 [103]. Uncontrolled expression of E6 leads to down-regulation of the *TP53* pathway and thus reduction of apoptosis contributing to cellular transformation to a malignant phenotype [110]. Deregulation of E7 leads to increased cell proliferation through many mechanisms, including degradation of the retinoblastoma protein [11], which controls the duration of the G1 phase of the cell cycle [43].

3.3 Hepatitis B virus

HBV DNA often integrates into hepatocellular carcinoma cells as infected hepatocytes regenerate [3, 12, 16, 25, 29, 70], but its necessity for carcinogenesis is uncertain [101]. Nonetheless, the number of HBV integrations in tumor cells is predictive of time of survival [126]. Individuals with >3 integrations typically have a decreased time of survival when compared to that of individuals with <3 integrations [126]. In addition, individuals with >3 integrations have an earlier age of onset of cancer and a lower incidence of cirrhosis, the key risk factor for development of hepatocellular carcinoma [126]. Therefore, the increase in HBV integrations may play a key role in hepatocellular carcinogenesis in younger individuals without cirrhosis [126].

HBV has been shown to integrate recurrently into seven human oncogenes and tumor suppressor genes: *TERT* [35, 86, 94, 106, 126], *ITPR1* (also known as *IP3R1*) [86, 94], *IRAK2* [94], *MAPK1* [86, 94], *MLL2* [86], *MLL4* [86, 106, 126], and *CCNE1* [126]. About 40% of the integrations are from a region of the HBV genome that contains the viral enhancer, the X gene, and the core gene [126]. Expression of the integrated HBV X protein and X antigen can transform liver cells to a malignant phenotype in culture [42, 137] and give rise to hepatocellular carcinoma in mice [49, 53, 137], in addition to altering host gene expression, thus increasing cell growth and survival [28, 29, 41, 98]. Some researchers speculate that integration of the viral enhancer could facilitate human-HBV fusion proteins, disruption of tumor suppressor genes, and dysregulation of other downstream genes [126].

3.4 Epstein-Barr virus

Burkitt’s lymphoma is a type of B-cell derived non-Hodgkin lymphoma that is highly aggressive and fast growing [81] and was linked to EBV in 1964 [83]. Burkitt’s lymphomas are classified as: endemic, sporadic, and immunodeficiency-related. Endemic Burkitt’s lymphoma is associated with endemic malaria in Africa [81, 87] where EBV is found integrated into 90% of endemic Burkitt’s lymphoma cases [97]. All types of Burkitt’s lymphoma, regardless of EBV infection, have a translocation involving the *MYC* oncogene, most commonly rearranged with an immunoglobulin chain gene [51, 127]. The *MYC*-immunoglobulin translocation constitutively activates cellular proliferation by putting an

immunoglobulin enhancer element in control of *MYC* expression, which is responsible for cell cycle control and contributes to carcinogenesis [127]. Sporadic cases of Burkitt's lymphoma typically have different translocation breakpoints compared to endemic cases of Burkitt's lymphoma, suggesting that EBV integration may play a role in generating the *MYC* translocation in endemic Burkitt's lymphoma [81].

3.5 Merkel cell polyomavirus

MCV is a common childhood infection that after extreme sun exposure can progress to Merkel cell carcinoma, a type of aggressive skin cancer [78]. MCV is clonally integrated into 80% of Merkel cell carcinomas at random genomic locations [30, 109]. MCV has a T-antigen protein complex known to target tumor suppressor proteins, such as retinoblastoma [121], and is solely expressed in tumor cells [120]. It has been suggested that in order for MCV to contribute to carcinogenesis there must be immunosuppression [83], then the virus must integrate into the host genome, and finally mutations must occur to the large T-antigen of the T-antigen protein complex, while maintaining an intact binding domain for retinoblastoma [83, 121]. This hypothesis has been backed by the knowledge that integrated MCV typically has a truncated C-terminal domain of the large T-antigen gene [64, 121]. The full length large T-antigen activates DNA damage responses leading to upregulation of *TP53* signaling, while the truncated large T-antigen does not have this capability instead triggering increased cell proliferation [64]. The large T-antigen truncation also causes loss of MCV-induced viral replication from the loss of C-terminal ATPase/helicase domain, which may allow the virus to evade the host immune system and contribute to carcinogenesis [121].

4. Microbial involvement in carcinogenesis

An estimated 15–20% of cancers worldwide are linked to infectious agents like viruses, parasites, and bacteria, resulting in 1.5 million cancer deaths worldwide in 2008 [20]. Given that there are 10X more bacterial cells in the human body than human cells [68], bacteria have a large potential to effect human health. As they relate to cancer, it is often difficult to establish causation. Microbial infections may contribute to an illness, or they may merely colonize as a result of microenvironment changes caused by the onset of disease. Yet clearly some bacteria are associated with cancer and some subset of those bacteria are thought to contribute to carcinogenesis.

Helicobacter pylori has been strongly associated with gastric carcinoma and gastric mucosa-associated lymphoid tissue (MALT) lymphoma [95]. *Schistosoma haematobium* has been associated with bladder carcinoma [31], *Salmonella typhi* with gallbladder cancer [24], *Chlamydia pneumonia* with lung cancer [66], and *Bacteroides fragilis*, *Streptococcus bovis*, *Escherichia coli* and *Fusobacterium* spp. with colorectal cancer [10, 15, 55, 56, 72, 105, 119]. Most of the bacteria described here have been shown to contribute to carcinogenesis through DNA damage initiated by increased inflammation [62] (Figure 1C). For example, *Helicobacter pylori* activates cyclooxygenase enzymes necessary for the formation of inflammatory prostaglandins. This increase in inflammation causes DNA damage by the release of mutagens like inducible nitric oxide synthase [95].

In addition, bacteria can secrete microbial proteins that affect cell signaling (Figure 1C). For example, enterotoxigenic *Bacteroides fragilis* secretes the *B. fragilis* toxin (BFT), a zinc-dependent metalloprotease toxin, which is highly related to diphtheria toxin [118]. BFT triggers cleavage of E-cadherin [135], which in turn activates β -catenin [136] and thus *Wnt* signaling, increasing colonic proliferation [71]. The *NF- κ B* pathway is also activated through interaction with BFT, causing the secretion of pro-inflammatory cytokines by colonic epithelial cells [118]. Therefore, like viruses, bacteria can contribute to carcinogenesis by expression of bacterial proteins and by causing increased inflammation [27, 96, 117] (Figure 1). The role of bacterial DNA integration in oncogenesis is less well studied.

5. Bacterial DNA integration as a putative carcinogen, challenges and future directions

5.1 Background

Viral integration is one example of lateral gene transfer (LGT), which describes the transfer of DNA between diverse organisms and is synonymous with horizontal gene transfer (HGT). LGT is most frequently described as occurring between different bacteria, but examples of LGT from bacteria to animals have been increasingly reported [21]. Bacteria-animal transfers frequently occur between endosymbionts and their host [21] and to specific nematodes and insects that parasitize plants [22]. The predominant limiting factor in the occurrence of LGT is the proximity between both organisms [7, 21], and in animals this means proximity to the germ cells [21]. One example of this are *Wolbachia* endosymbionts, germ line intracellular endosymbionts, which frequently transfer large portions of their genomes into their arthropod and nematode hosts [23]. However, genomic analysis of the arthropod mealybug genome reveals that multiple other germ line endosymbionts may also participate in such interactions [44].

5.2 Bacteria-human lateral gene transfer

If bacterial DNA can integrate into the germ cell genome of invertebrates and be inherited, as is the case with *Wolbachia* DNA, it seems feasible that bacterial DNA can become integrated into vertebrate somatic genomes. The bacterial type four secretion system of *Bartonella henselae* has been shown to transfer DNA into human cells in culture [115] as has the analogous system in *Agrobacterium tumefaciens* [58]. However, integration of microbial DNA into primary cancer cells has only been described recently [99]. It is known that human cells are protected against bacterial integration by both the innate immune response and the separation of germline cells from somatic cells [100]. For example, the human immune system recognizes bacterial DNA [37, 107] and mRNA [107] as pathogen-associated molecular patterns (PAMPs) and mounts an immune response. However, components of the immune system have not been identified that recognize all bacterial rRNA [90, 107]. Absence of such components would provide a gateway for integration of unrecognized bacterial rRNA into host genomes. In addition, immune reactions, while efficient, are also imperfect meaning that some nucleic acids may go undetected. Any integrations that result could then act as potential mutagens.

A recent study using sequencing data from The Cancer Genome Atlas reported evidence for integrations of *Pseudomonas-like* DNA into known tumor suppressor and proto-oncogenes in stomach adenocarcinoma samples [99]. The same report also described putative *Acinetobacter-like* DNA integrations into either the mitochondrial genome or nuclear mitochondrial transfers of acute myeloid leukemia samples [99]. In both cases, the putative integrations have homology to bacterial rRNA [99]. However, it does not demonstrate whether these integrations may be causative or if the cancer cells may merely be more permissive for such integrations. Previously, other interactions between microbes and human cells, such as chronic inflammation, have been implicated in bacteria-associated cancers, such as stomach adenocarcinoma [62]. However, the recent demonstration of integrations of *Pseudomonas-like* DNA into proto-oncogenes in a subset of TCGA stomach adenocarcinoma samples may indicate direct carcinogenic potential [99].

The integrations of *Pseudomonas-like* DNA into stomach adenocarcinoma samples were predominantly into five genes. Four of the five genes are proto-oncogenes, including *TMSB10*, *CEACAM5*, *CEACAM6*, and *CD74* [99]. All four of these genes are known to be up-regulated in gastric cancer [32, 39, 57, 89]. *CEACAM5* and *CEACAM6* are both carcinoembryonic antigen-related cell adhesion molecules, which are a type of immunoglobulin-related glycoproteins with varied functions including cell-cell recognition and regulation of cellular functions [57]. Specifically these molecules have been known to supervise tissue architecture [45] and serve as receptors for specific bacteria and viruses [57]. The CEA family of antigens is also one the most widely used tumor markers [9, 33]. Thymosin β 10 (*TMSB10*) is an actin monomer-sequestering protein that is involved in cell motility and migration through its ability to bind G-actin [138]. It can also alter angiogenesis, apoptosis, cell proliferation and many other cellular processes involved in cancer progression [124]. *CD74* is an important modulator of the immune response as it is a transmembrane glycoprotein that controls antigen presentation and is known to associate with major histocompatibility complex (MHC) class II [67]. *CD74* is also involved in various signaling pathways such as NF- κ B and AKT, which can both contribute to carcinogenesis [73, 125]. In the stomach adenocarcinoma samples containing *Pseudomonas-like* DNA insertions, the afore-mentioned genes possessed integrations of bacterial DNA in their 5'-untranslated region, which prompts speculation that the overexpression of these genes may have been triggered by the insertional mutagenesis of a repressor binding site [99] or a site governing transcript stability (Figure 2).

If bacterial integration is an important mechanism of oncogenesis, we expect that further studies in this area will reveal integrations analogous to those from mobile elements and viruses. Such integrations would include insertional inactivation of coding genes, and transfer of oncogenic proteins and peptides (e.g. the *B. fragilis* metalloprotease). Some phenomena found with mobile elements, like translocations, require multiple copies in the genome. Therefore, phenomena like translocations are not likely to involve bacterial integrations unless there are multiple integrations or duplications of the same piece of bacterial DNA.

Should bacterial LGT be an additional mechanism used by specific pathogenic bacteria to promote carcinogenesis, specific vaccines could be developed to prevent cancer

development, similar to that of HPV. The HPV vaccine is believed to decrease HPV-associated cancers by 70% [93] and serves as a great example of how a vaccine can be used to prevent cancer development caused by microbial insertional mutagenesis.

Future experiments examining bacteria-human LGT in cancer are necessary to validate these transfers and demonstrate their oncogenic potential. Effective validation methods include techniques used to validate mobile element and HPV integrations. For example, fluorescence *in situ* hybridization (FISH) is frequently used to visualize HPV integrations in human chromosomes [84, 129]. PCR amplification of the junction followed by sequencing is the standard experiment used to validate L1 insertions [46]. Southern blots are also frequently used to study mobile element break points [38, 77] and HPV integrations [26, 113] as they allow for detection of specific pieces of DNA. To evaluate the oncogenic potential of bacterial LGT in cancer, experimental approaches from the study of HPV could again be used. For example, expression constructs demonstrated that the presence of E6/E7 or the entire HPV genome in human cells could result in a cancerous phenotype [36, 48]. Finally, the recently developed CRISPR-Cas9 system, which enables efficient genome editing to modify specific genes in human cells and conduct *in situ* genetic screens [69], could be used to evaluate the impact of bacteria-human LGT.

5.3 Challenges of studying lateral gene transfer

Despite the evidence supporting the existence of LGT in animals, including humans, as well as the clinical benefits of identifying such an event, evaluation of such events has been challenging thus far. Most techniques used to identify foreign genetic material in a sample involve sequencing of PCR products, which carries the risk of contamination and detection of artifactual chimeric sequences. These chimeras would be formed in the ligation and/or PCR amplification during library construction [131, 132]. Unfortunately, it is hard to differentiate true LGT from chimerism, as sequence validation using PCR has the same risk of chimera formation [131, 132].

Another challenge in identifying LGT comes from assembling genome sequences. When genome assemblies are created they are often purged of reads that resemble bacterial sequences—either intentionally through foreign sequence screens or unintentionally through analysis of k-mers. In this situation it is possible that LGT from bacteria is present but never reflected in the final genome assembly [21]. Furthermore, modern mapping-based techniques that use BWA [63] and/or Bowtie [61] will miss bacterial reads if not included in the reference genome. Tumor heterogeneity is also a challenge for identifying LGT in cancer samples. The significance of sequencing data is based on the number of sequencing reads supporting that element, referred to as the sequence coverage [5]. If LGT does not occur early during clonal expansion it could be at such low levels that there may be no reads supporting it, or the coverage of the LGT may be so low that its validity is doubted.

5.4 Promising techniques to study lateral gene transfer

New sequencing technologies, such as Oxford Nanopore, will be extremely beneficial for investigation of LGT. As advertised, this new technology does not rely on amplifying DNA or on ligation of adapters [4, 114], which will significantly reduce the risk of generating

artifactual chimeric sequences in sequencing data [4]. Another benefit of this technology is that the starting sample is conserved during the steps leading to sequencing [4]. This improvement allows for verification of initial findings by validating them on the same starting sample, an important aspect for confirming rare variants in a heterogeneous sample [4]. When this technology becomes available it will allow for a more thorough examination of LGT in animals, potentially leading to the identification of disease-causing integrations in humans.

6. Concluding remarks

HPV, HBV, and EBV integrations into the human genome are a major part of carcinogenesis in their associated cancers [83]. *H. pylori* and *B. fragilis* also contribute considerably to their respective cancers, but through different mechanisms [31, 50, 119]. Microbial LGT in cancer genomes has not been studied extensively enough to definitively determine the frequency that such integrations occur and cause functionally significant mutations. With this in mind, it is important to consider all viruses and microbes present in tumors and thoroughly investigate all potential mechanisms for mutagenesis. More work needs to be done using up-and-coming sequencing technologies in order to critically evaluate the role of LGT from carcinogenic infectious agents. Only when LGT between bacteria and humans is given more careful consideration can we establish its relevance, or lack thereof, in human carcinogenesis.

Acknowledgments

This work was funded by the National Institutes of Health through the NIH Director's New Innovator Award Program (1-DP2-OD007372).

References

1. Abend JR, Jiang M, Imperiale MJ. BK virus and human cancer: innocent until proven guilty. *Semin Cancer Biol.* 2009; 19:252–260. [PubMed: 19505653]
2. Ansari KI, Mandal SS. Mixed lineage leukemia: roles in gene expression, hormone signaling and mRNA processing. *FEBS Journal.* 2010; 277:1790–1804. [PubMed: 20236313]
3. Arzumanyan A, Reis HM, Feitelson MA. Pathogenic mechanisms in HBV- and HCV-associated hepatocellular carcinoma. *Nat Rev Cancer.* 2013; 13:123–135. [PubMed: 23344543]
4. Astier Y, Braha O, Bayley H. Toward single molecule DNA sequencing: direct identification of ribonucleoside and deoxyribonucleoside 5'-monophosphates by using an engineered protein nanopore equipped with a molecular adapter. *J Am Chem Soc.* 2006; 128:1705–1710. [PubMed: 16448145]
5. Beal MA, C Glenn T, Somers CM. Whole genome sequencing for quantifying germline mutation frequency in humans and model species: cautious optimism. *Mutat Res.* 2012; 750:96–106. [PubMed: 22178956]
6. Bechtold V, Beard P, Raj K. Human papillomavirus type 16 E2 protein has no effect on transcription from episomal viral DNA. *J Virol.* 2003; 77:2021–2028. [PubMed: 12525636]
7. Beiko RG, Harlow TJ, Ragan MA. Highways of gene sharing in prokaryotes. *Proc Natl Acad Sci U S A.* 2005; 102:14332–14337. [PubMed: 16176988]
8. Benard A, van de Velde CJ, Lessard L, Putter H, Takeshima L, Kuppen PJ, Hoon DS. Epigenetic status of LINE-1 predicts clinical outcome in early-stage rectal cancer. *Br J Cancer.* 2013; 109:3073–3083. [PubMed: 24220694]

9. Benchimol S, Fuks A, Jothy S, Beauchemin N, Shiota K, Stanners CP. Carcinoembryonic antigen, a human tumor marker, functions as an intercellular adhesion molecule. *Cell*. 1989; 57:327–334. [PubMed: 2702691]
10. Biarc J, Nguyen IS, Pini A, Gosse F, Richert S, Thierse D, Van Dorsselaer A, Leize-Wagner E, Raul F, Klein JP, Scholler-Guinard M. Carcinogenic properties of proteins with proinflammatory activity from *Streptococcus infantarius* (formerly *S.bovis*). *Carcinogenesis*. 2004; 25:1477–1484. [PubMed: 14742316]
11. Boyer SN, Wazer DE, Band V. E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Res*. 1996; 56:4620–4624. [PubMed: 8840974]
12. Brechot C, Pourcel C, Louise A, Rain B, Tiollais P. Presence of integrated hepatitis B virus DNA sequences in cellular DNA of human hepatocellular carcinoma. *Nature*. 1980; 286:533–535. [PubMed: 6250074]
13. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, Kazazian HH Jr. Hot L1s account for the bulk of retrotransposition in the human population. *Proc Natl Acad Sci U S A*. 2003; 100:5280–5285. [PubMed: 12682288]
14. Carreira PE, Richardson SR, Faulkner GJ. L1 retrotransposons, cancer stem cells and oncogenesis. *FEBS J*. 2014; 281:63–73. [PubMed: 24286172]
15. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, Barnes R, Watson P, Allen-Verge E, Moore RA, Holt RA. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res*. 2012; 22:299–306. [PubMed: 22009989]
16. Chakraborty PR, Ruiz-Opazo N, Shouval D, Shafritz DA. Identification of integrated hepatitis B virus DNA and expression of viral RNA in an HBsAg-producing human hepatocellular carcinoma cell line. *Nature*. 1980; 286:531–533. [PubMed: 6250073]
17. Chiarle R, Zhang Y, Frock Richard L, Lewis Susanna M, Molinie B, Ho Y-J, Myers Darienne R, Choi Vivian W, Compagno M, Malkin Daniel J, Neuberger D, Monti S, Giallourakis Cosmas C, Gostissa M, Alt Frederick W. Genome-wide Translocation Sequencing Reveals Mechanisms of Chromosome Breaks and Rearrangements in B Cells. *Cell*. 2011; 147:107–119. [PubMed: 21962511]
18. Corden SA, Sant-Cassia LJ, Easton AJ, Morris AG. The integration of HPV-18 DNA in cervical carcinoma. *Mol Pathol*. 1999; 52:275–282. [PubMed: 10748877]
19. Das P, Thomas A, Mahantshetty U, Shrivastava SK, Deodhar K, Mulherkar R. HPV genotyping and site of viral integration in cervical cancers in Indian women. *PLoS One*. 2012; 7:e41012. [PubMed: 22815898]
20. de Martel C, Ferlay J, Franceschi S, Vignat J, Bray F, Forman D, Plummer M. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol*. 2012; 13:607–615. [PubMed: 22575588]
21. Dunning Hotopp JC. Horizontal gene transfer between bacteria and animals. *Trends Genet*. 2011; 27:157–163. [PubMed: 21334091]
22. Dunning Hotopp, JC. Lateral gene transfer in multicellular organisms. In: Gophna, U., editor. *Lateral gene transfer in evolution*. Springer Science; 2013. p. 161-179.
23. Dunning Hotopp JC, Clark ME, Oliveira DC, Foster JM, Fischer P, Munoz Torres MC, Giebel JD, Kumar N, Ishmael N, Wang S, Ingram J, Nene RV, Shepard J, Tomkins J, Richards S, Spiro DJ, Ghedin E, Slatko BE, Tettelin H, Werren JH. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science*. 2007; 317:1753–1756. [PubMed: 17761848]
24. Dutta U, Garg PK, Kumar R, Tandon RK. Typhoid carriers among patients with gallstones are at increased risk for carcinoma of the gallbladder. *Am J Gastroenterol*. 2000; 95:784–787. [PubMed: 10710075]
25. Edman JC, Gray P, Valenzuela P, Rall LB, Rutter WJ. Integration of hepatitis B virus sequences and their expression in a human hepatoma cell. *Nature*. 1980; 286:535–538. [PubMed: 6250075]
26. Edwards TG, Helmus MJ, Koeller K, Bashkin JK, Fisher C. Human papillomavirus episome stability is reduced by aphidicolin and controlled by DNA damage response pathways. *J Virol*. 2013; 87:3979–3989. [PubMed: 23365423]

27. Elinav E, Nowarski R, Thaiss CA, Hu B, Jin C, Flavell RA. Inflammation-induced cancer: crosstalk between tumours, immune cells and microorganisms. *Nat Rev Cancer*. 2013; 13:759–771. [PubMed: 24154716]
28. Feitelson MA, Duan LX. Hepatitis B virus X antigen in the pathogenesis of chronic infections and the development of hepatocellular carcinoma. *Am J Pathol*. 1997; 150:1141–1157. [PubMed: 9094970]
29. Feitelson MA, Lee J. Hepatitis B virus integration, fragile sites, and hepatocarcinogenesis. *Cancer Lett*. 2007; 252:157–170. [PubMed: 17188425]
30. Feng H, Shuda M, Chang Y, Moore PS. Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science*. 2008; 319:1096–1100. [PubMed: 18202256]
31. Fried B, Reddy A, Mayer D. Helminths in human carcinogenesis. *Cancer Lett*. 2011; 305:239–249. [PubMed: 20667649]
32. Gold DV, Stein R, Burton J, Goldenberg DM. Enhanced expression of CD74 in gastrointestinal cancers and benign tissues. *Int J Clin Exp Pathol*. 2010; 4:1–12. [PubMed: 21228923]
33. Goldstein MJ, Mitchell EP. Carcinoembryonic antigen in the staging and follow-up of patients with colorectal cancer. *Cancer Invest*. 2005; 23:338–351. [PubMed: 16100946]
34. Gostissa M, Alt FW, Chiarle R. Mechanisms that Promote and Suppress Chromosomal Translocations in Lymphocytes. *Annual Review of Immunology*. 2011; 29:319–350.
35. Gozuacik D, Murakami Y, Saigo K, Chami M, Mugnier C, Lagorce D, Okanoue T, Urashima T, Brechot C, Paterlini-Brechot P. Identification of human cancer-related genes by naturally occurring Hepatitis B Virus DNA tagging. *Oncogene*. 2001; 20:6233–6240. [PubMed: 11593432]
36. Gubanova E, Brown B, Ivanov SV, Helleday T, Mills GB, Yarbrough WG, Issaeva N. Downregulation of SMG-1 in HPV-positive head and neck squamous cell carcinoma due to promoter hypermethylation correlates with improved survival. *Clin Cancer Res*. 2012; 18:1257–1267. [PubMed: 22247495]
37. G, Hacker; Redecke, V.; Hacker, H. Activation of the immune system by bacterial CpG-DNA. *Immunology*. 2002; 105:245–251. [PubMed: 11918685]
38. Hadler HI, Devadas K, Mahalingam R. Selected nuclear LINE elements with mitochondrial-DNA-like inserts are more plentiful and mobile in tumor than in normal tissue of mouse and rat. *J Cell Biochem*. 1998; 68:100–109. [PubMed: 9407318]
39. Han SU, Kwak TH, Her KH, Cho YH, Choi C, Lee HJ, Hong S, Park YS, Kim YS, Kim TA, Kim SJ. *CEACAM5* and *CEACAM6* are major target genes for Smad3-mediated TGF-beta signaling. *Oncogene*. 2008; 27:675–683. [PubMed: 17653079]
40. Hancks DC, Kazazian HH Jr. Active human retrotransposons: variation and disease. *Curr Opin Genet Dev*. 2012; 22:191–203. [PubMed: 22406018]
41. Henkler F, Koshy R. Hepatitis B Virus transcriptional activators: Mechanisms and possible role in oncogenesis. *J Viral Hepatitis*. 1996; 3:109–121.
42. Hohne M, Schaefer S, Seifer M, Feitelson MA, Paul D, Gerlich WH. Malignant transformation of immortalized transgenic hepatocytes after transfection with hepatitis B virus DNA. *EMBO Journal*. 1990; 9:1137–1145. [PubMed: 2323335]
43. Hollingsworth RE Jr, Chen PL, Lee WH. Integration of cell cycle control with transcriptional regulation by the retinoblastoma protein. *Curr Opin Cell Biol*. 1993; 5:194–200. [PubMed: 8507491]
44. Husnik F, Nikoh N, Koga R, Ross L, Duncan RP, Fujie M, Tanaka M, Satoh N, Bachtrög D, Wilson AC, von Dohlen CD, Fukatsu T, McCutcheon JP. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell*. 2013; 153:1567–1578. [PubMed: 23791183]
45. Ilantzis C, Demarte L, Screaton RA, Stanners CP. Deregulated expression of the human tumor marker CEA and CEA family member *CEACAM6* disrupts tissue architecture and blocks colonocyte differentiation. *Neoplasia*. 2002; 4:151–163. [PubMed: 11896570]
46. Iskow RC, McCabe MT, Mills RE, Torene S, Pittard WS, Neuwald AF, Van Meir EG, Vertino PM, Devine SE. Natural mutagenesis of human genomes by endogenous retrotransposons. *Cell*. 2010; 141:1253–1261. [PubMed: 20603005]

47. Kazazian HH Jr, Wong C, Youssoufian H, Scott AF, Phillips DG, Antonarakis SE. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature*. 1988; 332:164–166. [PubMed: 2831458]
48. Khaiboullina SF, Morzunov SP, Hall MR, De Meirleir KL, Rizvanov AA, Lombardi VC. Human dendritic cells transfected with a human papilloma virus-18 construct display decreased mobility and upregulated cytokine production. *Int J Oncol*. 2013; 43:1701–1709. [PubMed: 23969559]
49. Kim CM, Koike K, Saito I, Miyamura T, Jay G. HBx gene of hepatitis B virus induces liver cancer in transgenic mice. *Nature*. 1991; 351:317–320. [PubMed: 2034275]
50. Kim SS, Ruiz VE, Carroll JD, Moss SF. *Helicobacter pylori* in the pathogenesis of gastric cancer and gastric lymphoma. *Cancer Lett*. 2011; 305:228–238. [PubMed: 20692762]
51. Klein G. The role of gene dosage and genetic transpositions in carcinogenesis. *Nature*. 1981; 294:313–318. [PubMed: 7312030]
52. Klein, Isaac A.; Resch, W.; Jankovic, M.; Oliveira, T.; Yamane, A.; Nakahashi, H.; Di Virgilio, M.; Bothmer, A.; Nussenzweig, A.; Robbiani, Davide F.; Casellas, R.; Michel, CNussenzweig. Translocation-Capture Sequencing Reveals the Extent and Nature of Chromosomal Rearrangements in B Lymphocytes. *Cell*. 2011; 147:95–106. [PubMed: 21962510]
53. Koike K, Moriya K, Iino S, Yotsuyanagi H, Endo Y, Miyamura T, Kurokawa K. High-level expression of hepatitis B virus HBx gene and hepatocarcinogenesis in transgenic mice. *Hepatology*. 1994; 19:810–819. [PubMed: 8138251]
54. Konkel MK, Batzer MA. A mobile threat to genome stability: The impact of non-LTR retrotransposons upon the human genome. *Semin Cancer Biol*. 2010; 20:211–221. [PubMed: 20307669]
55. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, Clancy TE, Chung DC, Lochhead P, Hold GL, El-Omar EM, Brenner D, Fuchs CS, Meyerson M, Garrett WS. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe*. 2013; 14:207–215. [PubMed: 23954159]
56. Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, Ojesina AI, Jung J, Bass AJ, Tabernero J, Baselga J, Liu C, Shivdasani RA, Ogino S, Birren BW, Huttenhower C, Garrett WS, Meyerson M. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res*. 2012; 22:292–298. [PubMed: 22009990]
57. Kuespert K, Pils S, Hauck CR. CEACAMs: their role in physiology and pathophysiology. *Curr Opin Cell Biol*. 2006; 18:565–571. [PubMed: 16919437]
58. Kunik T, Tzfira T, Kapulnik Y, Gafni Y, Dingwall C, Citovsky V. Genetic transformation of HeLa cells by *Agrobacterium*. *Proc Natl Acad Sci U S A*. 2001; 98:1871–1876. [PubMed: 11172043]
59. Kupperts R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene*. 2001; 20:5580–5594. [PubMed: 11607811]
60. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissole SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers

RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglu S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ. Initial sequencing and analysis of the human genome. *Nature*. 2001; 409:860–921. [PubMed: 11237011]

61. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009; 10:R25. [PubMed: 19261174]
62. Lax AJ, Thomas W. How bacteria could cause cancer: one step at a time. *Trends Microbiol*. 2002; 10:293–299. [PubMed: 12088666]
63. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009; 25:1754–1760. [PubMed: 19451168]
64. Li J, Wang X, Diaz J, Tsang SH, Buck CB, You J. Merkel cell polyomavirus large T antigen disrupts host genomic integrity and inhibits cellular proliferation. *J Virol*. 2013; 87:9173–9188. [PubMed: 23760247]
65. Li L, McVety S, Younan R, Liang P, Du Sart D, Gordon PH, Hutter P, Hogervorst FB, Chong G, Foulkes WD. Distinct patterns of germ-line deletions in *MLH1* and *MSH2*: the implication of *Alu* repetitive element in the genetic etiology of Lynch syndrome (HNPCC). *Hum Mutat*. 2006; 27:388. [PubMed: 16541406]
66. Littman AJ, White E, Jackson LA, Thornquist MD, Gaydos CA, Goodman GE, Vaughan TL. *Chlamydia pneumoniae* infection and risk of lung cancer. *Cancer Epidemiol Biomarkers Prev*. 2004; 13:1624–1630. [PubMed: 15466979]
67. Lotteau V, Teyton L, Peleraux A, Nilsson T, Karlsson L, Schmid SL, Quaranta V, Peterson PA. Intracellular transport of class II MHC molecules directed by invariant chain. *Nature*. 1990; 348:600–605. [PubMed: 2250716]
68. Luckey TD. Introduction to the ecology of the intestinal flora. *Am J Clin Nutr*. 1970; 23:1430–1432. [PubMed: 5475362]
69. Malina A, Mills JR, Cencic R, Yan Y, Fraser J, Schippers LM, Paquet M, Dostie J, Pelletier J. Repurposing CRISPR/Cas9 for in situ functional assays. *Genes Dev*. 2013; 27:2602–2614. [PubMed: 24298059]
70. Marion PL, Salazar FH, Alexander JJ, Robinson WS. State of hepatitis B viral DNA in a human hepatoma cell line. *J Virol*. 1980; 33:795–806. [PubMed: 6251250]
71. Markowitz SD, Bertagnolli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med*. 2009; 361:2449–2460. [PubMed: 20018966]
72. Martin HM, Campbell BJ, Hart CA, Mpofu C, Nayar M, Singh R, Englyst H, Williams HF, Rhodes JM. Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. *Gastroenterology*. 2004; 127:80–93. [PubMed: 15236175]
73. Matza D, Wolstein O, Dikstein R, Shachar I. Invariant chain induces B cell maturation by activating a TAFII105-NF- κ B-dependent transcription program. *J Biol Chem*. 2001; 276:27203–27206. [PubMed: 11371575]
74. McCabe MT, Low JA, Imperiale MJ, Day ML. Human polyomavirus BKV transcriptionally activates *DNA methyltransferase 1* through the pRb/E2F pathway. *Oncogene*. 2006; 25:2727–2735. [PubMed: 16547506]
75. Melsheimer P, Vinokurova S, Wentzensen N, Bastert G, von Knebel Doeberitz M. DNA aneuploidy and integration of human papillomavirus type 16 e6/e7 oncogenes in intraepithelial neoplasia and invasive squamous cell carcinoma of the cervix uteri. *Clin Cancer Res*. 2004; 10:3059–3063. [PubMed: 15131043]

76. Miki Y, Katagiri T, Kasumi F, Yoshimoto T, Nakamura Y. Mutation analysis in the *BRCA2* gene in primary breast cancers. *Nat Genet.* 1996; 13:245–247. [PubMed: 8640237]
77. Miki Y, Nishisho I, Horii A, Miyoshi Y, Utsunomiya J, Kinzler KW, Vogelstein B, Nakamura Y. Disruption of the *APC* gene by a retrotransposal insertion of L1 sequence in a colon cancer. *Cancer Res.* 1992; 52:643–645. [PubMed: 1310068]
78. Miller RW, Rabkin CS. Merkel cell carcinoma and melanoma: etiological similarities and differences. *Cancer Epidemiol Biomarkers Prev.* 1999; 8:153–158. [PubMed: 10067813]
79. Mills RE, Bennett EA, Iskow RC, Devine SE. Which transposable elements are active in the human genome? *Trends Genet.* 2007; 23:183–191. [PubMed: 17331616]
80. Mitchell RJ, Farrington SM, Dunlop MG, Campbell H. Mismatch repair genes *hMLH1* and *hMSH2* and colorectal cancer: a HuGE review. *Am J Epidemiol.* 2002; 156:885–902. [PubMed: 12419761]
81. Molyneux EM, Rochford R, Griffin B, Newton R, Jackson G, Menon G, Harrison CJ, Israels T, Bailey S. Burkitt's lymphoma. *Lancet.* 2012; 379:1234–1244. [PubMed: 22333947]
82. Montagna M, Santacatterina M, Torri A, Menin C, Zullato D, Chieco-Bianchi L, D'Andrea E. Identification of a 3 kb *Alu*-mediated *BRCA1* gene rearrangement in two breast/ovarian cancer families. *Oncogene.* 1999; 18:4160–4165. [PubMed: 10435598]
83. Moore PS, Chang Y. Why do viruses cause cancer? Highlights of the first century of human tumour virology. *Nat Rev Cancer.* 2010; 10:878–889. [PubMed: 21102637]
84. Mooren JJ, Gultekin SE, Straetmans JM, Haesevoets A, Peutz-Kootstra CJ, Huebbers CU, Dienes HP, Wieland U, Ramaekers FC, Kremer B, Speel EJ, Klussmann JP. P16(INK4A) immunostaining is a strong indicator for high-risk-HPV-associated oropharyngeal carcinomas and dysplasias, but is unreliable to predict low-risk-HPV-infection in head and neck papillomas and laryngeal dysplasias. *Int J Cancer.* 2014; 134:2108–2117. [PubMed: 24127203]
85. Mukherjee S, Mukhopadhyay A, Banerjee D, Chandak GR, Ray K. Molecular pathology of haemophilia B: identification of five novel mutations including a LINE 1 insertion in Indian patients. *Haemophilia.* 2004; 10:259–263. [PubMed: 15086324]
86. Murakami Y, Saigo K, Takashima H, Minami M, Okanoue T, Brechot C, Paterlini-Brechot P. Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas. *Gut.* 2005; 54:1162–1168. [PubMed: 16009689]
87. Niedobitek G, Agathangelou A, Rowe M, Jones EL, Jones DB, Turyaguma P, Oryema J, Wright DH, Young LS. Heterogeneous expression of Epstein-Barr virus latent proteins in endemic Burkitt's lymphoma. *Blood.* 1995; 86:659–665. [PubMed: 7605996]
88. Nussenzweig A, Nussenzweig MC. Origin of Chromosomal Translocations in Lymphoid Cancer. *Cell.* 2010; 141:27–38. [PubMed: 20371343]
89. Oien KA, Vass JK, Downie I, Fullarton G, Keith WN. Profiling, comparison and validation of gene expression in gastric carcinoma and normal stomach. *Oncogene.* 2003; 22:4287–4300. [PubMed: 12833151]
90. Oldenburg M, Kruger A, Ferstl R, Kaufmann A, Nees G, Sigmund A, Bathke B, Lauterbach H, Suter M, Dreher S, Koedel U, Akira S, Kawai T, Buer J, Wagner H, Bauer S, Hochrein H, Kirschning CJ. TLR13 recognizes bacterial 23S rRNA devoid of erythromycin resistance-forming modification. *Science.* 2012; 337:1111–1115. [PubMed: 22821982]
91. Oldridge M, Zackai EH, McDonald-McGinn DM, Iseki S, Morris-Kay GM, Twigg SRF, Johnson D, Wall SA, Jiang W, Theda C, Jabs EW, Wilkie AOM. De novo *Alu*-element insertions in *FGFR2* identify a distinct pathological basis for Apert syndrome. *Am J Hum Genet.* 1999; 64:446–461. [PubMed: 9973282]
92. Pace JK 2nd, Feschotte C. The evolutionary history of human DNA transposons: evidence for intense activity in the primate lineage. *Genome Res.* 2007; 17:422–432. [PubMed: 17339369]
93. Parkin DM. The global health burden of infection-associated cancers in the year 2002. *Int J Cancer.* 2006; 118:3030–3044. [PubMed: 16404738]
94. Paterlini-Brechot P, Saigo K, Murakami Y, Chami M, Gozuacik D, Mugnier C, Lagorce D, Brechot C. Hepatitis B virus-related insertional mutagenesis occurs frequently in human liver cancers and recurrently targets human *telomerase* gene. *Oncogene.* 2003; 22:3911–3916. [PubMed: 12813464]

95. Peek RM Jr, Blaser MJ. *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. *Nat Rev Cancer*. 2002; 2:28–37. [PubMed: 11902583]
96. Plottel CS, Blaser MJ. Microbiome and malignancy. *Cell Host Microbe*. 2011; 10:324–335. [PubMed: 22018233]
97. Purtilo DT, Manolov G, Manolova Y, Harada S, Lipscomb H, Tatsumi E. Role of Epstein-Barr virus in the etiology of Burkitt's lymphoma. *IARC Sci Publ*. 1985:231–247. [PubMed: 2998991]
98. Rawat S, Clippinger A, Bouchard M. Modulation of apoptotic signaling by the hepatitis B virus X protein. *Viruses*. 2012; 4:2945–2972. [PubMed: 23202511]
99. Riley DR, Sieber KB, Robinson KM, White JR, Ganesan A, Nourbakhsh S, Dunning Hotopp JC. Bacteria-human somatic cell lateral gene transfer is enriched in cancer samples. *PLoS Comput Biol*. 2013; 9:e1003107. [PubMed: 23840181]
100. Robinson KM, Sieber KB, Dunning Hotopp JC. A review of bacteria-animal lateral gene transfer may inform our understanding of diseases like cancer. *PLoS Genet*. 2013; 9:e1003877. [PubMed: 24146634]
101. Robinson WS, Klote L, Aoki N. Hepadnaviruses in cirrhotic liver and hepatocellular carcinoma. *J Med Virol*. 1990; 31:18–32. [PubMed: 2165515]
102. Rocha PP, Skok JA. The origin of recurrent translocations in recombining lymphocytes: a balance between break frequency and nuclear proximity. *Current Opinion in Cell Biology*. 2013; 25:365–371. [PubMed: 23478218]
103. Romanczuk H, Howley PM. Disruption of either the E1 or the E2 regulatory gene of human papillomavirus type 16 increases viral immortalization capacity. *Proc Natl Acad Sci U S A*. 1992; 89:3159–3163. [PubMed: 1313584]
104. Roy AM, Carroll ML, Nguyen SV, Salem AH, Oldridge M, Wilkie AO, Batzer MA, Deininger PL. Potential gene conversion and source genes for recently integrated Alu elements. *Genome Res*. 2000; 10:1485–1495. [PubMed: 11042148]
105. Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/beta-catenin signaling via its FadA adhesin. *Cell Host Microbe*. 2013; 14:195–206. [PubMed: 23954158]
106. Saigo K, Yoshida K, Ikeda R, Sakamoto Y, Murakami Y, Urashima T, Asano T, Kenmochi T, Inoue I. Integration of hepatitis B virus DNA into the myeloid/lymphoid or *mixed-lineage leukemia (MLL4)* gene and rearrangements of *MLL4* in human hepatocellular carcinoma. *Hum Mutat*. 2008; 29:703–708. [PubMed: 18320596]
107. Sander LE, Davis MJ, Boekschoten MV, Amsen D, Dascher CC, Ryffel B, Swanson JA, Muller M, Blander JM. Detection of prokaryotic mRNA signifies microbial viability and promotes immunity. *Nature*. 2011; 474:385–389. [PubMed: 21602824]
108. Sarid R, Gao SJ. Viruses and human cancer: from detection to causality. *Cancer Lett*. 2011; 305:218–227. [PubMed: 20971551]
109. Sastre-Garau X, Peter M, Avril MF, Laude H, Couturier J, Rozenberg F, Almeida A, Boitier F, Carloti A, Couturaud B, Dupin N. Merkel cell carcinoma of the skin: pathological and molecular evidence for a causative role of MCV in oncogenesis. *J Pathol*. 2009; 218:48–56. [PubMed: 19291712]
110. Scheffner M, Huibregtse JM, Vierstra RD, Howley PM. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell*. 1993; 75:495–505. [PubMed: 8221889]
111. Schichman SA, Caligiuri MA, Gu Y, Strout MP, Canaani E, Bloomfield CD, Croce CM. *ALL-1* partial duplication in acute leukemia. *Proc Natl Acad Sci U S A*. 1994; 91:6236–6239. [PubMed: 8016145]
112. Schichman SA, Caligiuri MA, Strout MP, Carter SL, Gu Y, Canaani E, Bloomfield CD, Croce CM. *ALL-1* tandem duplication in acute myeloid leukemia with a normal karyotype involves homologous recombination between *Alu* elements. *Cancer Res*. 1994; 54:4277–4280. [PubMed: 8044771]
113. Schmitz M, Driesch C, Beer-Grondke K, Jansen L, Runnebaum IB, Durst M. Loss of gene function as a consequence of human papillomavirus DNA integration. *Int J Cancer*. 2012; 131:E593–E602. [PubMed: 22262398]

114. Schneider GF, Dekker C. DNA sequencing with nanopores. *Nat Biotechnol.* 2012; 30:326–328. [PubMed: 22491281]
115. Schröder G, Schuelein R, Quebatte M, Dehio C. Conjugative DNA transfer into human cells by the VirB/VirD4 type IV secretion system of the bacterial pathogen *Bartonella henselae*. *Proc Natl Acad Sci U S A.* 2011
116. Schulz TF, Boshoff CH, Weiss RA. HIV infection and neoplasia. *Lancet.* 1996; 348:587–591. [PubMed: 8774574]
117. Schwabe RF, C Jobin. The microbiome and cancer. *Nat Rev Cancer.* 2013; 13:800–812. [PubMed: 24132111]
118. Sears CL. Enterotoxigenic *Bacteroides fragilis*: a Rogue among Symbiotes. *Clin Microbiol Rev.* 2009; 22:349–369. [PubMed: 19366918]
119. Sears CL, Pardoll DM. Perspective: alpha-bugs, their microbial partners, and the link to colon cancer. *J Infect Dis.* 2011; 203:306–311. [PubMed: 21208921]
120. Shuda M, Arora R, Kwun HJ, Feng H, Sarid R, Fernandez-Figueras MT, Tolstov Y, Gjoerup O, Mansukhani MM, Swerdlow SH, Chaudhary PM, Kirkwood JM, Nalesnik MA, Kant JA, Weiss LM, Moore PS, Chang Y. Human Merkel cell polyomavirus infection I. MCV T antigen expression in Merkel cell carcinoma, lymphoid tissues and lymphoid tumors. *Int J Cancer.* 2009; 125:1243–1249. [PubMed: 19499546]
121. Shuda M, Feng H, Kwun HJ, Rosen ST, Gjoerup O, Moore PS, Chang Y. T antigen mutations are a human tumor-specific signature for Merkel cell polyomavirus. *Proc Natl Acad Sci U S A.* 2008; 105:16272–16277. [PubMed: 18812503]
122. Society, AC. Cancer Facts & Figures 2013. Atlanta: American Cancer Society; 2013.
123. Solyom S, Ewing AD, Rahrmann EP, Doucet T, Nelson HH, Burns MB, Harris RS, Sigmon DF, Casella A, Erlanger B, Wheelan S, Upton KR, Shukla R, Faulkner GJ, Largaespada DA, Kazazian HH Jr. Extensive somatic L1 retrotransposition in colorectal tumors. *Genome Res.* 2012; 22:2328–2338. [PubMed: 22968929]
124. Sribenja S, Li M, Wongkham S, Wongkham C, Yao Q, Chen C. Advances in thymosin beta10 research: differential expression, molecular mechanisms, and clinical implications in cancer and other conditions. *Cancer Invest.* 2009; 27:1016–1022. [PubMed: 19909017]
125. Stein R, Mattes MJ, Cardillo TM, Hansen HJ, Chang C-H, Burton J, Govindan S, Goldenberg DM. CD74: A New Candidate Target for the Immunotherapy of B-Cell Neoplasms. *Clinical Cancer Research.* 2007; 13:5556s–5563s. [PubMed: 17875789]
126. Sung WK, Zheng H, Li S, Chen R, Liu X, Li Y, Lee NP, Lee WH, Ariyaratne PN, Tennakoon C, Mulawadi FH, Wong KF, Liu AM, Poon RT, Fan ST, Chan KL, Gong Z, Hu Y, Lin Z, Wang G, Zhang Q, Barber TD, Chou WC, Aggarwal A, Hao K, Zhou W, Zhang C, Hardwick J, Buser C, Xu J, Kan Z, Dai H, Mao M, Reinhard C, Wang J, Luk JM. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nat Genet.* 2012; 44:765–769. [PubMed: 22634754]
127. Taub R, Kirsch I, Morton C, Lenoir G, Swan D, Tronick S, Aaronson S, Leder P. Translocation of the *c-myc* gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc Natl Acad Sci U S A.* 1982; 79:7837–7841. [PubMed: 6818551]
128. Teugels E, De Brakeleer S, Goelen G, Lissens W, Sermijn E, De Grève J. De novo *Alu* element insertions targeted to a sequence common to the *BRCA1* and *BRCA2* genes. *Human Mutation.* 2005; 26:284–284. [PubMed: 16088935]
129. Theelen W, Litjens RJ, Vinokurova S, Haesevoets A, Reijans M, Simons G, Smedts F, Herrington CS, Ramaekers FC, von Knebel Doeberitz M, Speel EJ, Hopman AH. Human papillomavirus multiplex ligation-dependent probe amplification assay for the assessment of viral load, integration, and gain of telomerase-related genes in cervical malignancies. *Hum Pathol.* 2013; 44:2410–2418. [PubMed: 23968641]
130. Van Tine BA, Kappes JC, Banerjee NS, Knops J, Lai L, Steenbergen RD, Meijer CL, Snijders PJ, Chatis P, Broker TR, Moen PT Jr, Chow LT. Clonal selection for transcriptionally active viral oncogenes during progression to cancer. *J Virol.* 2004; 78:11172–11186. [PubMed: 15452237]

131. Wang GC, Wang Y. Frequency of formation of chimeric molecules as a consequence of PCR coamplification of 16S rRNA genes from mixed bacterial genomes. *Applied and Environmental Microbiology*. 1997; 63:4645–4650. [PubMed: 9406382]
132. Wang GCY, Wang Y. The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. *Microbiology*. 1996; 142:1107–1114. [PubMed: 8704952]
133. Wentzensen N, Vinokurova S, von Knebel Doeberitz M. Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer Res*. 2004; 64:3878–3884. [PubMed: 15172997]
134. Whitman SP, Liu S, Vukosavljevic T, Rush LJ, Yu L, Liu C, Klisovic MI, Maharry K, Guimond M, Strout MP, Becknell B, Dorrance A, Klisovic RB, Plass C, Bloomfield CD, Marcucci G, Caligiuri MA. The *MLL* partial tandem duplication: evidence for recessive gain-of-function in acute myeloid leukemia identifies a novel patient subgroup for molecular-targeted therapy. *Blood*. 2005; 106:345–352. [PubMed: 15774615]
135. Wu S, Lim KC, Huang J, Saidi RF, Sears CL. *Bacteroides fragilis* enterotoxin cleaves the zonula adherens protein, E-cadherin. *Proc Natl Acad Sci U S A*. 1998; 95:14979–14984. [PubMed: 9844001]
136. Wu S, Morin PJ, Maouyo D, Sears CL. *Bacteroides fragilis* enterotoxin induces *c-Myc* expression and cellular proliferation. *Gastroenterology*. 2003; 124:392–400. [PubMed: 12557145]
137. You X, Liu F, Zhang T, Li Y, Ye L, Zhang X. Hepatitis B virus X protein upregulates oncogene Rab18 to result in the dysregulation of lipogenesis and proliferation of hepatoma cells. *Carcinogenesis*. 2013; 34:1644–1652. [PubMed: 23471881]
138. Yu FX, Lin SC, Morrison-Bogorad M, Atkinson MA, Yin HL. Thymosin beta 10 and thymosin beta 4 are both actin monomer sequestering proteins. *J Biol Chem*. 1993; 268:502–509. [PubMed: 8416954]

Mobile elements can cause cancer by inserting into cancer-related genes.

Viruses contribute to carcinogenesis by mutating and deregulating human genes.

Microbes induce chronic inflammation as well as toxins that are linked to cancer.

Viral and mobile element integration allude to microbial integration in the cancer genome.

New technologies may facilitate detecting microbial integrations in the human genome.

Figure 1. Carcinogenesis by mobile element movement, viral integrations, and microbial infections

(A) Mobile genetic elements, such as retrotransposons, can play a role in cancer mutagenesis. They can either move to a new location in the genome or recombine with similar mobile elements and cause genomic instability. This vulnerability often contributes to cancer progression. (B) Viruses use three broad mechanisms to promote carcinogenesis. Viruses can integrate into the human genome inducing genomic instability (purple arrow on the left). The deregulation of signaling pathways induced from these chromosomal disruptions transform normal cells into cells with cancerous phenotypes. Regardless of life

cycle stage, viruses can express proteins that alter normal cell cycle patterns by increasing proliferation and decreasing apoptosis (purple arrow on the right). Viruses can also alter inflammatory responses that cause DNA damage and contribute to cancer development (gray arrow), a mechanism that is not discussed in this review. (C) Microbial infections can lead to carcinogenesis through multiple different pathways. It is important to note that these pathways can occur in conjunction with each other or individually. Microbes can first cause irritation and inflammation that leads to DNA damage and thus cancer (gray arrow). Or bacteria can secrete microbial proteins that alter host-signaling pathways, leading to increased cell proliferation and decreased apoptosis, the classic characteristic of cancer (blue arrow on the right). While multiple examples of viral integrations that progress to cancer have been described, evidence for microbial integration into the human genome has only been reported recently.

Figure 2. Illustration of stomach adenocarcinoma bacterial integration sites in relation to other gene features

Four of the five genes implicated in *Pseudomonas-like* DNA integrations in stomach adenocarcinoma samples are depicted. Their bacterial integration sites are shown as well as the exons, introns, direction of translation, and untranslated regions.