

## Minireview

# Infectious causes of cancer and their detection

Lucy Dalton-Griffin\* and Paul Kellam\*†

Addresses: \*Department of Infection, University College London, Cleveland Street, London W1T 4JF, UK. †The Wellcome Trust Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

Correspondence: Paul Kellam. Email: p.kellam@ucl.ac.uk

### Abstract

Molecular techniques for identifying pathogens associated with cancer continue to be developed, including one reported recently in *BMC Medical Genomics*. Identifying a causal infectious agent helps in understanding the biology of these cancers and can lead ultimately to the development of antimicrobial drugs and vaccines for their treatment and prevention.

In 1911, Peyton Rous used cell-free filtered extract of a chicken sarcoma to establish an association between cancer and an infectious agent - the Rous sarcoma virus. Almost 100 years later, developments in the techniques used to detect microbial genomes and investigate their biological properties have led to a definitive role for viral, bacterial and parasitic infection in human carcinogenesis. Recent estimates indicate that the average proportion of malignancies worldwide that could be avoided in the absence of an infectious agent is 17.8%, with this figure being higher in developing countries, at 26.3% [1].

The development of cancer is a complex multistage process that, as outlined by Hanahan and Weinberg [2], can be grouped into “six essential alterations to the cell physiology” (Figure 1). These changes most likely occur by a progressive, almost evolutionary mechanism of successive genetic changes. The main mechanisms by which chronic infection promotes cancer do not usually involve direct mutagenesis, but instead are due to the complex interactions that occur between host and pathogen. In order for viruses to replicate and persist, they often promote cell survival, drive cellular proliferation, and evade the immune system. Together, these processes, especially if present over a long period of time, can lead to tumorigenesis, often by influencing the same pathways that are involved in the development of cancer in the absence of infection (Figure 1).

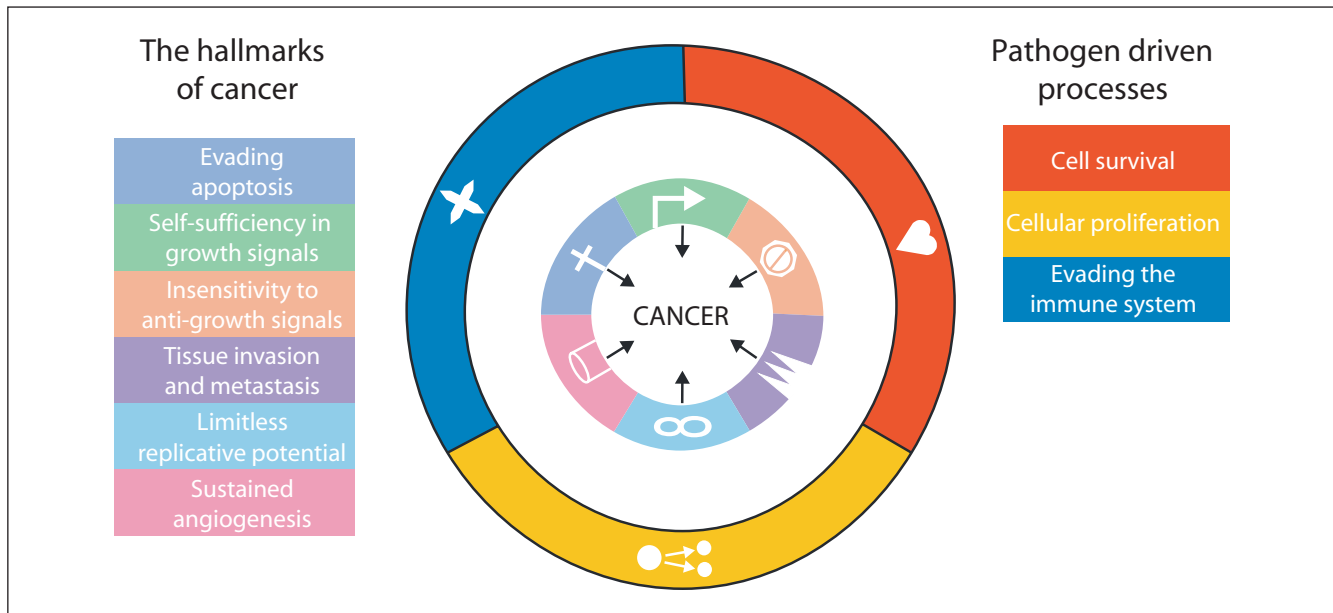
Indeed, several of the key mediators of pathways and networks proposed by Hanahan and Weinberg were discovered through the study of viruses. Many oncogenes, for example *Ras* and *Myc*, were identified originally in cancer-causing retroviruses. Similarly, the study of DNA

virus proteins, such as SV40 large T-antigen, was instrumental in discovering tumor suppressor genes such as p53. Understanding the biology of oncogenic infections will most likely continue to inform cancer biology in general. One challenge is that of identifying potentially oncogenic agents and proving their causal connection with cancer. A recent paper in *BMC Medical Genomics* by Duncan *et al.* [3] describes the new computational technique of digital karyotyping microbe identification (DK-MICROBE), and its application to identifying pathogens. Their paper further exemplifies both the potential of molecular and bioinformatics methods for identifying pathogen DNA in tumor samples, but underlines the necessity of establishing a causal association rather than just a pathogen presence.

### How infectious agents cause cancer

The major mechanisms by which infectious agents can promote and maintain tumor formation can be divided broadly into three main categories (Figure 2). The first is the induction of chronic inflammation as a result of a continuing immune response to a persistent infection. This occurs, for example, in the case of hepatitis C virus (HCV), associated with liver cancer, which continually replicates in the liver, setting up a chronic state of inflammation there. Similarly, the blood fluke *Schistosoma haematobium* and the Gram-negative bacterium *Helicobacter pylori* can both directly contribute to cancer formation through persistence within the host causing chronic inflammation [4]. *H. pylori* is a good example of this category, and was classified by the World Health Organization as a class 1 carcinogen in 1994. There is a high prevalence of persistent infection with *H. pylori*: worldwide, 75% of people are infected, with prevalence being higher in sub-Saharan Africa, where *H. pylori* is associated with 63.4% of all stomach cancers [1]. However, the fact that not all people infected with *H. pylori* develop gastric cancer clearly shows that the infectious agent is a risk factor, but that other environmental and genetic influences are involved in cancer formation.

Second, oncogenesis can occur through virus-induced transformation. This is due to the persistence of the viral genome in a latent form in an infected cell, either without



**Figure 1**

Continued persistent infection by a pathogen (outer circle) requires host-cell survival (red), host-cell proliferation (yellow), and evasion of the immune system by the pathogen (blue). These pathogen-driven processes are achieved via various mechanisms that interfere with normal cell physiology and are outlined in Figure 2. Alterations in these normally highly regulated pathways can lead to transforming events that have been described as the 'hallmarks of cancer' (inner circle) [2]. Accumulation of such events can lead to cancer development. Cancer is not, however, an outcome that has been specifically selected by evolution to aid pathogen survival. Rather, it is more likely an unfortunate coincidence of pathogen capabilities selected to enable successful infection. Therefore, certain infections may not necessarily cause the infected individual to develop cancer, but may be an associated risk factor (Figure adapted from [2]).

replication, as with Epstein-Barr virus (EBV), which infects B lymphocytes, or through integration of the viral genome into a host-cell chromosome, as with human papillomavirus (HPV), the cause of cervical cancer. EBV is frequently detected in childhood Burkitt's lymphoma, post-transplant B-cell lymphomas, non-Hodgkin's lymphoma, Hodgkin's disease and nasopharyngeal carcinoma [1]. The transforming capability of this virus is exemplified further by its ability to transform resting B cells *in vitro* at high efficiency to obtain stable proliferating lymphoblastoid cell lines. This process is driven by EBV-encoded latent proteins that directly promote cell growth and survival - for example, lymphocyte membrane-associated protein-1 (LMP-1).

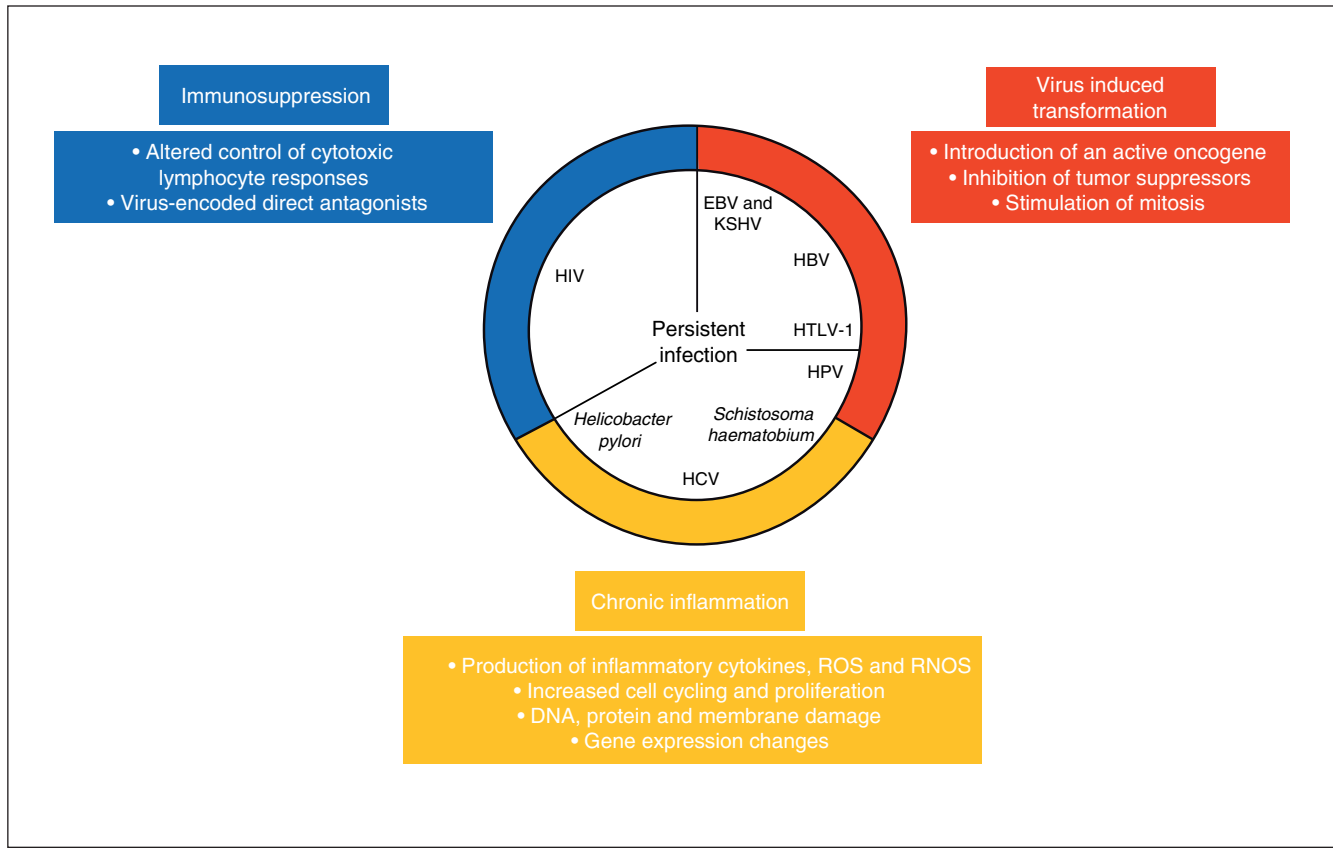
The third mechanism is the chronic suppression of the immune system by the infectious agent, such as the immunodeficiency (AIDS) caused by HIV infection. The presence of natural mechanisms of immunosurveillance for cancer cells, which in the case of an infectious etiology will also involve immune mechanisms that routinely control the infection, suggests why pathogens with oncogenic potential do not rapidly cause malignancy. A compromised immune system can result in an increased incidence of infection-driven tumors by weakening the immune control. Such an increase is seen, for example, in transplant

patients, who are being treated with immunosuppressants, or in individuals with AIDS [5].

Pathogens associated with cancer exemplify many of these mechanisms; persistent infection involves evading the immune response as well as chronic inflammation, which even in the immune-competent leads to chronic cell proliferation and a greater risk of oncogenic transformation. However, many non-oncogenic pathogens are equally adept at these processes, indicating that other factors must be involved. For example, the risk of an infectious agent causing cancer may also depend on the cell type infected, as certain cell lineages could be more 'prone' to transformation than others. For example, the increased prevalence of lymphomas and leukemias in children and young adults suggests that lymphocytes are more susceptible to transformation.

### Detecting infectious agents in cancer

'The harder you look the more you find' seems to be true for infectious agents in disease. The ability of some agents to remain latent, as well as the existence of new and emerging infections [6], make detection and proving causality challenging. An infectious agent may trigger the initial events of oncogenesis but be absent in the final tumor, which adds timing of detection to the problem.



**Figure 2**

Infectious agents can contribute to malignant transformation by several mechanisms. These can be broadly divided into: chronic inflammation, which drives abnormal levels of cell proliferation (yellow); direct virus-induced transformation of infected cells, leading to increased cell survival (red); and immunosuppression, which allows the pathogen to evade the immune system and persist (blue). The colour coding is maintained from Figure 1. Chronic inflammation leads to the production of inflammatory cytokines as well as reactive oxygen and nitrogen oxide species (ROS and RNOS) by phagocytes at the site of infection, which can lead to DNA damage as well as cellular damage and increased cell cycling. Virus-induced transformation is caused by the actions of pathogen-encoded oncogenic proteins as well as by integration into the host genome (HPV). The transforming events outlined in this figure do not necessarily lead directly to cancer formation; for example, despite encoding similar proteins, other infectious agents do not cause cancer. The fact that some pathogens have evolved to persist without causing tumorigenesis also highlights that persistence is maybe a prerequisite for, but is on its own insufficient for, oncogenesis in humans. Immune evasion mechanisms include control of the adaptive and innate immune system, allowing avoidance of tumor surveillance. EBV, Epstein-Barr virus; HBV, human hepatitis virus B; HCV, hepatitis virus C; HIV, human immunodeficiency virus; HPV, human papillomavirus; HTLV-1, human T-lymphotropic virus 1; KSHV, Kaposi sarcoma-associated herpesvirus.

However, once the causal agent has been unequivocally found, development of a preventive treatment can be relatively rapid, as has been the case for cervical cancer. Human papilloma viruses, especially types HPV16 and HPV18, are now firmly associated with cervical cancer after their discovery in 1983, and this has led to the development and widespread use of an HPV vaccine within 25 years - a timescale not dissimilar to that of drug discovery. Similarly, recovery of immune function in AIDS by inhibiting HIV replication by antiretroviral therapy can lead to a regression of Kaposi's sarcoma, an endothelial tumor caused by the herpesvirus KSHV, and a decrease in incidence of other AIDS-related cancers.

Treating infection is therefore a valuable addition to antitumor therapy if the infectious agent can be identified. Today, two main techniques are used to detect microbial genomes in disease, one based on hunting for acknowledged candidates and the other on removing (either physically or computationally) known human sequences to reveal any foreign nucleic acid. PCR and microarray-based strategies are limited by a finite number of probes and sequences available but can be very sensitive. Subtractive methods include representational difference analysis (RDA), which was used to detect KSHV in Kaposi's sarcoma in 1994. They do, however, require isogenic controls, which are not always readily available.

Table 1

**Molecular guidelines for establishing microbial causation of disease**

Criteria	Causal relationship
Putative pathogen genome is present in most cases of disease	Microbial nucleic acid should be found preferentially in diseased sites in combination with anatomic, histologic, chemical or clinical evidence of pathology and not in areas lacking the pathology
Only diseased tissue should harbor putative pathogen genome	Fewer, or no, copy numbers of pathogen-associated nucleic acid sequences should occur in non-diseased host or tissue
Disease resolution should be accompanied by a reduction in copy number of pathogen genome	Disease resolution perhaps due to effective clinical treatment should lead to undetectable or reduced pathogen-associated nucleic acid. Any relapse in disease should see an increase in copy number
Microbial sequence may be detected before disease or may correlate with disease severity	A causal relationship can be more strongly inferred when pathogen-associated nucleic acid is present before disease onset and copy number correlates with disease severity
The nature of the microbial organism associated by detection of its nucleic acid should be consistent with known biological characteristics of that group of organisms	When phenotypes such as pathology, microbial morphology and clinical features are predicted by sequence-based phylogeny the meaningfulness of the detected sequence can be enhanced
Microbe-associated sequences detected in disease tissue should be corroborated at the cellular level	<i>In situ</i> hybridization of microbial sequences in an area of tissue pathology (or where microorganisms are thought to be located) should be attempted
Molecular evidence should be reproducible	Any sequence-based evidence for microbial causation must be replicated

Koch's postulates for proving a causal connection between a particular infectious agent and a disease cannot be applied to many human diseases as it would be unethical to experimentally infect humans with a potentially lethal infectious agent. The development of molecular diagnostic technology has enabled the criteria for causality summarized here to be drawn up [10].

One new approach to finding pathogen genomes is that of Duncan *et al.* [3], which applies computational subtraction to digital karyotyping to hunt for virus genomes in several primary colorectal cancer samples and metastases as well as in normal tissue. The technique of DK-MICROBE described by Duncan *et al.* [3] aims to circumvent limitations on detection imposed by the different mechanisms by which pathogens contribute to disease. In DK-MICROBE, genomic DNA from the tumor is digested enzymatically into fragments of less than 10 kb in size that are processed to yield 21-bp tags for amplification, concatenation and sequencing. Human sequences are computationally removed; the remaining unidentified 'pathogen' tags are then studied further. However, DK-MICROBE in its present form can only detect the genomes of DNA viruses. Duncan *et al.* [3] were able to detect the human herpesvirus 6 (HHV6) genome in samples from tumor tissue, but the fact that they also identified the viral DNA in healthy tissues well illustrates the difficulties of causally associating a particular virus with a particular cancer.

Another subtractive method known as digital transcript subtraction (DTS) attempts to identify exogenous pathogen transcripts via high-throughput sequencing, and can thus potentially identify the presence of RNA and DNA viruses [7]. This method involves developing a long serial analysis of gene expression (L-SAGE) library from the tumor cells by quantitatively joining 21-bp tags composed of cDNA copied from the 3' end of mRNAs. It can therefore detect all transcripts that are expressed in the tumor. As all

human tumor viruses to date express part of their genome in the transformed cells, this has proved effective in virus discovery. The pioneers of this technique, Yuan Chang and Patrick Moore, have validated DTS by identifying sequences from KSHV in the primary effusion lymphoma cell line BCBL-1 (Feng *et al.* [7]). More recently, DTS has been used to identify a new polyomavirus in an uncommon but aggressive human skin cancer, Merkel cell carcinoma (MCC) [8]. A fusion transcript between an unknown virus T-antigen and a human receptor tyrosine kinase was detected. The new virus was named Merkel cell polyomavirus (MCV) and was detected in 80% of MCC tumors and also in 16% of normal skin biopsies. In 75% of the MCV-related MCCs viral DNA was integrated in a clonal pattern, suggesting a potential mechanism for transformation. MCC occurs predominantly in the elderly and immunosuppressed, two of the key features that indicate an infectious etiological agent.

### Where to find associations?

The question of where to look for other cancer-causing infectious agents is partly answered by the example of MCC. Cancers with increased incidence in HIV-infected individuals [5] or in transplant recipients are an ideal place to look. One example is squamous cell conjunctival carcinoma (SCCC), which has emerged with the AIDS epidemic and is common in parts of sub-Saharan Africa. Papillomaviruses have been implicated, but the evidence from PCR-based studies and serology testing is controversial. Feng *et al.* [7] attempted to identify viral transcripts

using DTS, finding 21 candidate sequences that did not align with the human genome. However, further analysis revealed that they were all most likely of human origin [7]. This does not rule out an infectious aetiology for SCCC, however. DTS is a powerful tool, but it is important to realize that its sensitivity is governed by the depth of sequencing, that is, the number of sequence reads analysed by DTS or DK MICROBE, which may have to be increased in tumor samples, in which the ratio of viral mRNA to human mRNA is low. Nevertheless, the combination of carefully selected tumors with deep sequencing and computational identification of non-human sequences is set to uncover more tumor viruses.

Detection of a microbial genome in a tumor does not prove causation, but the classical Koch's postulates for proving the microbial cause of a disease cannot be invoked in this setting. One cannot, for example, infect a person with a suspected causal agent and then wait to see if they develop cancer. Modified criteria for infection, namely Hill's nine criteria outlined in 1965, provide an evaluation method that allows weight to be given to causative factors [9]. These were further developed by Fredricks and Relman to take into account molecular evidence (Table 1) [10]. These rules, when applied to the detection of pathogen genomes, can serve as a basis for the association of an infectious agent with a particular disease. New techniques will undoubtedly lead to the identification of new and existing infectious agents in disease tissues, including cancer. However, care must be taken not to overemphasize association without proper proof of causation.

## Acknowledgements

We wish to thank Robin A Weiss for reading a draft manuscript and making helpful comments.

## References

1. Parkin DM: **The global health burden of infection-associated cancers in the year 2002.** *Int J Cancer* 2006, **118**:3030-3044.
2. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **100**:57-70.
3. Duncan CG, Leary RJ, Lin JC, Cummins J, Di C, Schaefer CF, Wang TL, Riggins GJ, Edwards J, Bigner D, Kopelovich L, Vogelstein B, Kinzler KW, Velculescu VE, Yan H: **Identification of microbial DNA in human cancer.** *BMC Med Genomics* 2009, **2**:22.
4. Lax AJ, Thomas W: **How bacteria could cause cancer: one step at a time.** *Trends Microbiol* 2002, **10**:293-299.
5. Stein L, Urban MI, O'Connell D, Yu XQ, Beral V, Newton R, Ruff P, Donde B, Hale M, Patel M, Sitas F: **The spectrum of human immunodeficiency virus-associated cancers in a South African black population: results from a case-control study, 1995-2004.** *Int J Cancer* 2008, **122**:2260-2265.
6. Morens DM, Folkers GK, Fauci AS: **The challenge of emerging and re-emerging infectious diseases.** *Nature* 2004, **430**:242-249.
7. Feng H, Taylor JL, Benos PV, Newton R, Waddell K, Lucas SB, Chang Y, Moore PS: **Human transcriptome subtraction by using short sequence tags to search for tumor viruses in conjunctival carcinoma.** *J Virol* 2007, **81**:11332-11340.
8. Feng H, Shuda M, Chang Y, Moore PS: **Clonal integration of a polyomavirus in human Merkel cell carcinoma.** *Science* 2008, **319**:1096-1100.
9. Hill AB: **The environment and disease: association or causation?** *Proc R Soc Med* 1965, **58**:295-300.
10. Fredericks DN, Relman DA: **Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates.** *Clin Microbiol Rev* 1996, **9**:18-33.

Published: 11 August 2009

doi:10.1186/jbiol68

© 2009 BioMed Central Ltd