Article title: A reversion to the bacterial world - bacterial DNA triggering ancient response mechanisms as a potential cause of cancer in humans.
Authors: Paul Sharma[1]
Affiliations: Costain plc (London University)[1]
Orcid ids: 0000-0002-5864-893X[1]
Contact e-mail: paul@paulsharma.com

Funder: No funding received.
DOI: 10.14293/S2199-1006.1.SOR-.PPHUCBW.v1
Preprint first posted online: 18 January 2023
Keywords: lateral gene transfer, horizontal gene transfer, serial atavism model, cancer, hypothesis, review.

A reversion to the bacterial world - bacterial DNA triggering ancient response mechanisms as a potential cause of cancer in humans.

Authors: Paul Sharma*

Keywords: lateral gene transfer, horizontal gene transfer, serial atavism model, cancer, hypothesis, review.

Abstract
Lateral gene transfer, also known as horizontal gene transfer, is a common occurrence within the bacterial world and is a less frequent, but nevertheless, an ongoing feature of the eukaryotic environment. Several studies, reviewed in this paper, have reported an increased likelihood of cancer in humans when there is a higher preponderance of bacterial lateral gene transferred material within the cell. Separately, it has been suggested that cancer in humans may be linked to an atavistic response to DNA damage in the cell by bacterially inherited DNA. It is proposed that a mechanism for cancer in humans is the triggering of ancient bacterially-derived response mechanisms by bacterial genetic material entering eukaryotic cells – a reversion to the bacterial world. The reversion to the bacterial world hypothesis can be tested by large scale and longitudinal sampling.

Affiliations: University College, London

*Corresponding author: paul.sharma@costain.com, Costain Group PLC, Vanwell Business Park, 1 Vanwall Rd, Maidenhead SL6 4UB, United Kingdom.

Introduction

This paper reviews the evidence of bacterial DNA entering the cell via lateral gene transfer (LGT) and bacterial DNA integration into the human genome. Separately, there is the hypothesis that the genes corrupted in cancer which lead to loss of regulatory function are those related to the eukaryotic human cell's bacterial past. It is proposed that cancer is caused by LGT from bacteria entering the human cell, triggering bacterial-like responses.

There is increasing evidence that bacterial DNA exists in human cancers through genome sequencing, PCR amplification, 16S RNA detection and KEGG (Kyoto Encyclopaedia of Genes and Genomes)/TCGA (the cancer genome atlas) database comparison. Evidence of bacteria located within human cancer cells was analysed by Nejman et al. who reported that the bacteria within tumours varied according to tumour type. The analysis of the tumour microbiome, studying 1526 tumours and their adjacent normal tissues across seven cancer types, found that each tumour type has a distinct microbiome composition. The intratumour bacteria were mostly intracellular and were present in both cancer and immune cells. [1]

Yangyanqiu and Shuwen [2] in their recent comprehensive review of bacterial DNA linkage in carcinogenesis have suggested three sources of bacterial DNA in cancers: bacterial DNA transferred laterally by bacteria near human cells, bacterial DNA remaining in human blood after bacterial infection, and bacterial DNA passed down from progenitors.

This paper focusses on the first of these mechanisms, lateral gene transfer, also known as horizonal gene transfer [3], as a means for bacterial DNA to enter eukaryotic cells. Generally, LGT describes the processes involving the transfer of genes across cell boundaries when cells are near each other, involving three processes – transformation, transduction, and conjugation [3].

There is general acceptance that LGT was the key to the development of the eukaryotic cell by the nuclear acquisition of genes from mitochondria and chloroplasts, which originated from α-protobacteria and cyanobacteria respectively [4]. While LGT is widespread in the bacterial world driving genetic diversity and increasing fitness, a key example being the transfer of antibiotic resistance genes between bacteria, there are also many examples of LGT from bacteria to eukaryotes [3]. DNA transfer from bacteria to human may be much more prevalent than observed at present, as units smaller than a gene entering the cell via LGT are not recognised using phylogenetic analysis [2, 4].

LGT and cancer linkage

The linkage between viruses and cancer has been well established with key examples including cervical cancer caused by human papillomavirus DNA integration and liver cancer caused by hepatitis B virus DNA integration [2]. It was noted by Riley et al., that the large number of microbes within the human body (approximately 10x times the numbers of cells in the body) [5] could give rise to bacterial DNA integrations (BDIs) [3].

The concept that bacterial genes can be integrated into the human genome has become more prevalent since the work of Riley et al. of around ten years ago. The reader is recommended the review of this area by Yangyanqiu and Shuwen (October 2022) as an accompaniment to this paper [2].

While the majority of BDIs should be neutral in impact, it is conceivable that some BDIs could integrate into areas that promote carcinogenesis. There is evidence of bacterial-human LGT up to 210x higher [5] within cancer tumour samples following analyses of data from The Cancer Genome Atlas (TCGA), with the highest matches found for acute myeloid leukaemia, potentially via gene fragments into the mitochondrial genome, followed by stomach adenocarcinoma [5]. When regions on the human genome coverage were examined by Sieber et al. [3], a pattern emerged of integration in the mitochondria for myeloid leukaemia and five genes in stomach adenocarcinoma. In the case of stomach adenocarcinoma, four of those five genes have previously been shown to be implicated in cancer [6–9].

Further to the review of Yangyanqiu and Shuwen [2] it is noted that the work in bacterial DNA and carcinogenesis also includes Al-Abbas and Abdul-Ridha (2017), who examined 96 patients with hematological malignancies and from 26 healthy individuals, with 72 (75%) bacterial 16S rDNA genes detected in the DNA of patients with malignancies, but none in the 26 healthy individuals. The majority of bacterial DNA integrations in the human DNA arose from *Stenotrophomonas maltophilia*, *Massilia timonae* and *Methylobacterium lusitanum* [10].

In addition, Sidhoo, Rosales and Lee (2017) reported that a 17-bp homologous sequence in *Marinobacter sp*. Hb8, *Rhodococcus fascians* D188, *Rhodococcus sp*. PBTS2, *Micrococcus luteus* strain trpE16 and *M. luteus* NCTC 2665 integrated into the genome of an eosinophilic leukaemia patient [11]. Arrieta et al. (2022) evaluated the presence of the IS6110 *Mycobacterim tuberculosis* (Mtb) transposon on the primary tumour DNA of eighty-eight patients through end-point PCR, with results showing that 40.9% (36/88) had a positive end-point PCR for the IS6110 transposon [12]. Shotgun sequencing from two samples identified traces of MTb genomes present in lung adenocarcinomas tumour tissue, suggesting that similar Mtb strains could be infecting both patients.

As noted by Yangyanqiu and Shuwen [2], Akimova et al. (2022), analysed paired-end RNA sequencing data of 45 chronic lymphocytic leukaemia (CLL) patients and 9 healthy donors, looking for bacterial DNA integrations [13]. Their CLL samples demonstrated bacterial DNA integrations were approximately two-fold compared to normal samples. It was noted that LGT seen in gene CD74 has also been linked to stomach adenocarcinomas. In addition, Akimova et al. saw a high integration rate of LGT detected in class switched memory B cells, indicating that LGT may increase with the lifespan of a cell and with the activity of the DNA repair machinery [13].

In addition to these studies, an investigation by Borchmann (2021) using datasets obtained from multiple International Cancer Genome Consortium studies on more than 3000 whole genome sequencing datasets to investigate links between viruses, bacteria, and cancer found that 218 species-level taxa could be identified in tumour tissue, and of these 27 taxa were likely cancer-linked [14]. While Borchmann observed bacterial DNA within cancer cells, the only taxa where there was DNA integration into the host genome was in the case of *Hepatitis B* [14].

Finally, Bakhti and Latifi-Navid (2021) noted that gut microbiota is in close relation with humans and markedly influence gastric cancer with chronic *Helicobacter pylori* infection being a critical risk factor for gastric cancer [15], although carcinogenic mechanism could be inflammation caused by a chronic, mostly subclinical infection [14]. *Fusobacterium nucleatum* can also be found throughout the cancerous tissue of colorectal cancer at much higher levels than in the tissue of benign adenomas or healthy colon mucosa [14]. It has been proposed that there could be cross-interactions that occur both before and after tumorigenesis and during cancer progression [16].

Cancer as ancestral life form

Bacterial DNA can induce changes such as strand breaks and genomic instability, enabling DNA to enter the human genome via template sequence insertion [17], the integration of DNA to repair DNA double stranded breaks. This mutagenic form of DNA repair may play a role in genetic disease, exon shuffling, and mammalian evolution [18].

Work undertaken using the COSMIC and DAVID databases to examine the areas within the genome which are related to cancer mutations, noted that recessive cancer genes were clustered around the functional areas of cell cycle control and DNA repair, and that cancer cells in humans have a pattern of damage around DNA repair areas [19]. Genes involved in

DNA repair in humans are older than the average and are related to the genes in bacteria that respond to stressors. Rosenberg et al. noted that bacteria respond to DNA breakages by switching to a lower fidelity but faster system which repairs the double-strand with a higher rate of mutations, which is an adaptive response to stress [20].

It has been posited that cancer is a response to a series of atavistic reversions, most recently the Serial Atavism Model [21], whereby cancer hallmarks roughly correlate with the age of the relevant genes, and is a sequence of atavistic reversions, rather than a multicellular-to-unicellular switch. It has been noted that that the genes responsible for cellular cooperation in multicellular organisms (e.g., signalling, adhesion, angiogenesis and migration) are those genes that are corrupted in cancer and lead to loss of regulatory function [21, 22], in effect the cell reverting back to the bacterial world, where there is not the cooperation between cells as seen in the eukaryotic world. It has been posited that human cells become cancerous as these genes activate a repair response initiating a high rate of mutation as their reaction to stress [23] replicating their function in bacteria.

Work undertaken using phylostratigraphy from seven solid cancers, obtained from comparisons between the gene age and gene expression level in RNA sequencing data from TCGA, has confirmed that genes present in unicellular organisms were strongly upregulated, whereas genes of metazoan origin were primarily inactivated [24, 25]. Trigos et at. noted rewiring of the coupling between the gene networks that control unicellular processes from those that control multicellular processes [24]. *In vitro* analysis of the emergence of doxorubicin resistance in multiple myeloma cancer cells support the above thesis [26].

The hypothesis

There is evidence of LGT coinciding with cancer following the analysis of tumour samples and databases, as outlined above. Separately it has been reported that in cancer cells there is damage to the areas where there are genes responsible for cellular cooperation in multicellular organisms - the serial atavistic model.

The central hypothesis of this paper is to draw together these two stands of work, LGT incursions into eukaryotic cells and the serial atavism model into a single framework, a reversion to the bacterial world. The central tenet is that cells in humans become cancerous as bacterial DNA within the human cell arising from LGT initiates bacterial DNA integrations and double-strand breaks, which in turn activates an atavistic bacterially-derived repair mechanism creating high rates of mutation, which can lead to cancer.

Hypothesis testing

The direct work in cancer samples as outlined above [1, 2, 3, 5, 10, 11, 12, 13] has indicated the potential linkage of LGT in cancer, and studies could be expanded to cover larger datasets and include longitudinal tracking of larger patient cohorts. While most LGT insertions should be neutral in impact, large scale studies using healthy and non-healthy donors would indicate if there were hot spots on the genome that cause cancer, if these hotspots changed over time, and if these areas were linked to ancestral genes as suggested by the proponents of the serial atavism model [21].

Given the decreasing costs of human genome sequencing, it will be possible to compare the DNA content of individual cells in frequent contact with bacteria with those from more sterile portions of the body to understand the overall extent of LGT occurrence [4]. Further work in the area will clarify if cancer is "a return to an ancestral life form" and "that besides being a disease, cancer is also a form of life," as noted by Doru Paul [27].

The same testing programme outlined for this hypothesis would also be relevant for the Dong and Xing alternative suggestion regarding the bacterial origin of cancer cells [28], whereby cancer cells are new single-celled eukaryotes, formed by hybridising acquired eukaryotic DNA and bacterial DNA.

Discussion

It is known that viruses promote human cancer, with most viral-related human cancers caused by human papillomaviruses (HPV), hepatitis B and C viruses, and Epstein-Barr virus, with viruses causing around 10-15% of cancers [29]. The main carcinogenic mechanisms for viral carcinogenesis are thought to be via viral integration into and disruption of the host genome and expression of oncogenic viral proteins [30].

While the linkage between virus promotion and human cancer is accepted, it is hypothesised that there is also a linkage between bacterial DNA and cancer. A higher level of LGT seen in human cancer cells than non-cancerous cells has been ascertained, although the causality is uncertain – it could be that LGT is increased in cancer cells rather than vice versa. It is likely that that this would be resolved via the large-scale longitudinal testing outlined above.

DNA fragments from bacteria are not detectable at present using phylogenetic analysis, and therefore LGT could be much more prevalent than is viewed at present [4], it is noted is that the highest cancer rate correlated with LGT is found where there is a greater level of biome, namely in the stomach [5]. In addition, LGT appears to be associated with mitochondria and myeloid leukaemia and in these circumstances bacterial DNA would need to pass through only the outer cellular wall and not enter the nucleus.

Bacterial DNA can enter the nucleus of host cells though nuclear pores via the bifunctional and bidirectional hydrophilic nucleocytoplasmic exchange channel. One potential LGT transformation mechanism is for DNA fragments to enter cells is via dissipative diffusion, especially for smaller gene fragments, as many biological processes take place at near-equilibrium thermodynamic levels [31]. The stochastic nature of DNA diffusion could be a factor determining the occurrence of cancer, with the other factors being the level of exposure to bacterial DNA fragments, and the condition of cellular walls. This could be a contributory factor behind cancers exhibiting genomic alterations and mutations across a range of scales from single nucleotides to entire chromosomes [32]. Recent studies have shown that microbial nucleic acid not only exist in plasma and tumour cells of cancer patients, and can also be detected in the peripheral blood of healthy individuals [16].

Conclusion

This paper proposes that bacterial DNA insertion via LGT could be a cause of human cancer. As BDIs are rare triggers at the individual cellular level, incursions are most likely to be

observed in samples taken from patients with cancer, which has been demonstrated in smaller-scale studies. It is suggested that these studies are expanded in scale to include longitudinal screening.

On the more general question of LGT in humans, as screening costs decline, it will be possible to sample individual cells to understand whether LGT and small fragment bacterial DNA diffusion occurs in humans, and at what scale and frequency. It will also be possible to understand the level of BDI occurrence, whether there is a link to LGT and DNA diffusion, and it is proposed that this might be an area of future research.

Conflict of interest
The author declares there is no conflict of interest. No funding has been received for this paper.

References

[1] Nejman B et al. The human tumor microbiome is composed of tumor type–specific intracellular bacteria. Science, Vol 368, No 6494, pp 973-980. May 2020. DOI: 10.1126/science.aay9189

[2] Yangyanqiu W and Shuwen H. Bacterial DNA involvement in carcinogenesis. Front. Cell. Infect. Microbiol. 12:996778, 2022. doi: 10.3389/fcimb.2022.996778

[3] Sieber KB, Bromley RB and Dunning Hotopp JC. Lateral gene transfer between prokaryotes and eukaryotes. Experimental Cell Research 358: 2017; 421-426.

[4] Dunning Hotopp JC. Horizontal gene transfer between bacteria and animals. Trends in Genetics 2011: 27(4): 157-163.

[5] Riley DR, Sieber KB, Robinson KM, White JR, et al. Bacteria-human somatic lateral gene transfer is enriched in cancer samples. PLOS Computational Biology 2013; 9 Issue 6.

[6] 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.

[7] Kuespert K, Pils S, Hauck CR. CEACAMs: their role in physiology and pathophysiology. Current Opinion in Cell Biology 2006; 18: 565–571.

[8] Han SU, Kwak TH, Her KH, Cho YH, Choi C, et al. CEACAM5 and CEACAM6 are major target genes for Smad3-mediated TGF-beta signaling. Oncogene 2008; 27: 675–683.

[9] Gold DV, Stein R, Burton J, Goldenberg DM. Enhanced expression of CD74 in gastrointestinal cancers and benign tissues. Int J Clin Exp Pathol 2011; 4 (1): 1–12.

[10] Al-Abbas MJ, Lamyaa A, and Abdul-Ridha LA. Practical Genetic Study Confirming and Identifying the Horizontal Gene Transfer (HGT) of Bacterial DNA (16S rDNA Gene) Integrated

in the DNA of Cancer Patients. Merit Research Journal of Medicine and Medical Sciences 2017: (ISSN: 2354-323X) Vol. 5(2) pp. 060-071.

[11] Sidhoo S, Rosales J. and Ki-Young L. Integration of a bacterial gene sequence into a chronic eosinophilic leukemia patient's genome as part of a fusion gene linker. Biomarker Research: 2017; 5:20.

[12] Arrieta O, Molina-Romero C, Cornejo-Granados F, et al. Clinical and pathological characteristics associated with the presence of the IS6110 *Mycobacterim tuberculosis* transposon in neoplastic cells from non-small cell lung cancer patients. Scientific reports 2022; 12.1: 2210–2210.

[13] Akimova E, Gassner FJ, Griel R, et al. Detecting Bacterial–Human Lateral Gene Transfer in Chronic Lymphocytic Leukemia. International journal of molecular sciences 2022; 23.3: 1094.

[14] Borchmann S. An atlas of the tissue and blood metagenome in cancer reveals novel links between bacteria, viruses and cancer. Microbiome 2021; 9:94.

[15] Bakhti SZ, and Latifi-Navid S. Interplay and cooperation of *Helicobacter pylori* and gut microbiota in gastric carcinogenesis. BMC microbiology 2021: 21:258.

[16] You L, Zhou J, Xin Z, et al. Novel directions of precision oncology: circulating microbial DNA emerging in cancer-microbiome areas. Precision Clinical Medicine, 2022; 5: pbac005

[17] Onozawa M, Zhang Z, Kim YJ, Goldberg L, Varga T, Bergsagel PL, Kuehla WM, and Aplan PD. Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. PNAS 2014; 111(21): 7729–7734.

[18] Morrish TA, Gilbert N, Myers JS, Vincent BJ, Stamato TD, Taccioli GE, Batzer MA and Moran JV. DNA repair mediated by endonuclease-independent LINE-1 retrotransposition. Nature Genetics 2002; 31: 159-165.

[19] Cisneros L, Bussey KJ, Orr AJ, Miočević M, Lineweaver CH and Davies P. Ancient genes establish stress-induced mutation as a hallmark of cancer. PLoS ONE 2017; 12(4): e0176258

[20] Rosenberg SM, Shee C, Frisch RL, and Hastings PJ. Stress-induced mutation via DNA breaks in *Escherichia coli*: A molecular mechanism with implications for evolution and medicine. Bioessays 2012; 34: 885–892

[21] Lineweaver CH, Bussey KJ, Anneke C. Blackburn AC and Davies PCW. Cancer progression as a sequence of atavistic reversions. BioEssays 2021; 43: 2000305.

[22] Domazet-Lošo T and Tautz D. Phylostratigraphic tracking of cancer genes suggest a link to the emergence of multicellularity in metazoa. BMC Biology 2010, 8:66.

[23] Davies PCW and Lineweaver CH. Cancer tumors as Metazoa 1.0: tapping genes of ancient ancestors. Physical Biology 2011; 8 015001.

[24] Bussey KJ., Cisneros LH, Lineweaver CW, Davies PCW. Ancestral gene regulatory networks drive cancer. Proc. Natl. Acad. Sci. Unit. States Am. 2017; 114 (24), 6160e6162.

[25] Trigos A, Pearson RB, Papenfuss AT, Goode DL. Altered interactions between unicellular and multicellular genes drive hallmarks of transformation in a diverse range of solid tumors. Proc. Natl. Acad. Sci. Unit. States Am. 2017; 114 (24),

[26] Wu A, Zhang Q, Lambert G, et al. Ancient Hot and Cold Genes and Chemotherapy Resistance Emergence. Proceedings of the National Academy of Sciences – PNAS 2015; 112.33: 10467–10472.

[27] Doru P. 2021. Cancer as a form of life: Musings of the cancer and evolution symposium. Progress in biophysics and molecular biology 2021; 165: 120–139.

[28] Dong L, and Xing X. Cancer cells arise from bacteria. Cancer Cell International 2018; 18.1: 205–9.

[29] Bouza E, Jiménez MM, Alemany L, et al. Overview of virus and cancer relationships. Position paper. Revista española de quimioterapia: 2021; 34.6: 525–555.

[30]. Moore PS and Chang Y. Why do viruses cause cancer? Highlights of the first century of human tumour virology. Nat Rev Cancer 2010; 10: 878–889.

[31] Crooks GE. Entropy production fluctuation theorem and the nonequilibrium work regulation for free-energy differences. Phys. Rev. 1999; E 60, 2721–2726.

[32] Bussey KJ and Davies PCW. Reverting to single-cell biology: The predictions of the atavism theory-of cancer. Progress in Biophysics and Molecular Biology 2021; 165 49-55.