Microbiome genomics for cancer prediction

Although cancer genomics is a powerful tool to understand cancer and develop diagnostic tools, the contribution of the microbiome in cancer diagnosis and clinical assessment is much less studied. Elinav, Greten and colleagues provide their respective views on how studying cancer metagenomes could facilitate identification, diagnosis and staging of different tumor types.

Lorenz Adlung, Eran Elinav, Tim F. Greten and Firouzeh Korangy

From the bench: Lorenz Adlung and Eran Elinav

The advent of advanced 'omics' technologies, including whole-genome sequencing and single-cell transcriptomics, has led to the realization that cancer—rather than arising solely from single genetic events—is fueled by a complex network of contributing factors. An increased capacity to acquire, decode and utilize patients' individual genetic information has been instrumental in improving the accuracy and effectiveness of cancer diagnosis and treatment, but these approaches still insufficiently explain tumor initiation and progression at the level of the individual. In addition to the human genome, the microbial genome, termed the microbiome, may integrate another level of functional complexity that might contribute to personalized cancer prevention, diagnosis and therapy. Indeed, microbiome alterations have been recently correlated to cancer development, progression and treatment response in mice and in humans[1–3].

Writing in *Nature*, Knight and colleagues[4] add to this mounting evidence by proposing a novel method that utilizes microbial DNA from tissue and blood samples, enabling the discrimination of cancer from healthy tissue, the distinction between various types of cancer, and even the identification of different stages within certain tumor types. To systematically characterize the cancer-associated microbiome, the researchers re-analyzed published data from The Cancer Genome Atlas (TCGA) across 33 cancer types for microbial reads within more than 18,000 tissue samples. One of the biggest challenges of their approach was to achieve a robust detection of sequencing reads assigned to bacteria, archaea or viruses and to discern them from contaminants or technical errors. The authors achieved this by carefully benchmarking their classification, normalization, decontamination and batch-correction methods. Using the resulting dataset to subsequently train machine-learning models, they identified specific cancer-associated microbial signatures. An encouraging result pointing to the potential validity of their approach was obtained in gastrointestinal cancer samples in which an over-representation of *Fusobacterium* spp. aligned with previous reports[5]. The same retrospective approach, coupled with in silico–based sample decontamination methods, was applied to blood plasma samples from an alternative cohort encompassing 100 cancer patients and 69 control subjects. This allowed the authors to distinguish between cancer and control samples as well as between different cancer types (for instance prostate, lung and skin) by plasma-derived microbial profiling.

This elegant proof-of-concept study shows that blood plasma, an easily accessible material, may be suitable for the detection of characteristic microbial signatures contributing to a more sensitive and accurate cancer diagnosis. Of note, high sensitivity is required to reliably detect microbial DNA in blood, as it represents a low biomass microbiome[6]. Equally challenging may be differentiation between living and degraded components of microbes in this context. These will likely represent formidable challenges to be tackled in follow-up works.

In addition to introducing a potential new cancer diagnosis, patient stratification and treatment personalization pipeline, these results may also help establish a causal link between the presence of individual microbes or microbial consortia and physiological or pathological properties of human cancer. Indeed, the gut microbiome, the extra-intestinal microbiomes and even the recently suggested cancer microbiome[7] may contribute to the dynamic interplay between a host's genetics, immunity, the environment and cancer-related clinical and histopathological properties. An emerging challenge in metagenomic analysis is to establish causation in these otherwise associative studies. The machine-learning links established in this elegant work may be followed up mechanistically in future work, for example, aiming to elucidate how the microbial milieu participates in tumor development, growth and metastasis and to study tumor surveillance, antitumor immunity and treatment responses. Such mechanistic insights may translate into a better understanding of individual variability in cancer susceptibility, metastasis potential as well as personalized response to cancer treatment, eventually leading to the development of data-driven patient-specific therapeutics and prognostic markers.

In this context, one potential link from the microbiome to host cancer pathology is represented by metabolites produced, modified or degraded by bacteria. These potentially bioactive molecules could be locally produced at the tumor or generated distally, from which they could then systemically influx into the host and influence tumor-related processes. Metabolite profiles and their effector functions may vary between individuals harboring different microbiomes or exposed to different diets, medications or other environmental factors. Combinations of these microbiome-derived molecules could be either beneficial or detrimental in the context of cancer or its therapy response, and may offer a 'patient-specific chemical signature' that in combination with the analysis of microbial ecosystem composition by analyzing a patient's metagenome, might provide a new facet for personalized intervention. Future inclusion of systemic metabolite signature may further enhance the predictive capacity of the pipeline developed by Knight and colleagues[4] by integrating the current taxonomy assignments with functional microbiome signatures of metabolic pathways.

The notion that compositional microbial signatures can be indicative of individual host physiology is increasingly established in the context of nutrition[8], cardiometabolic diseases[9] and other 'multifactorial' human disorders. Knight and colleagues[4] contribute to this global effort by establishing a pipeline for the robust detection of

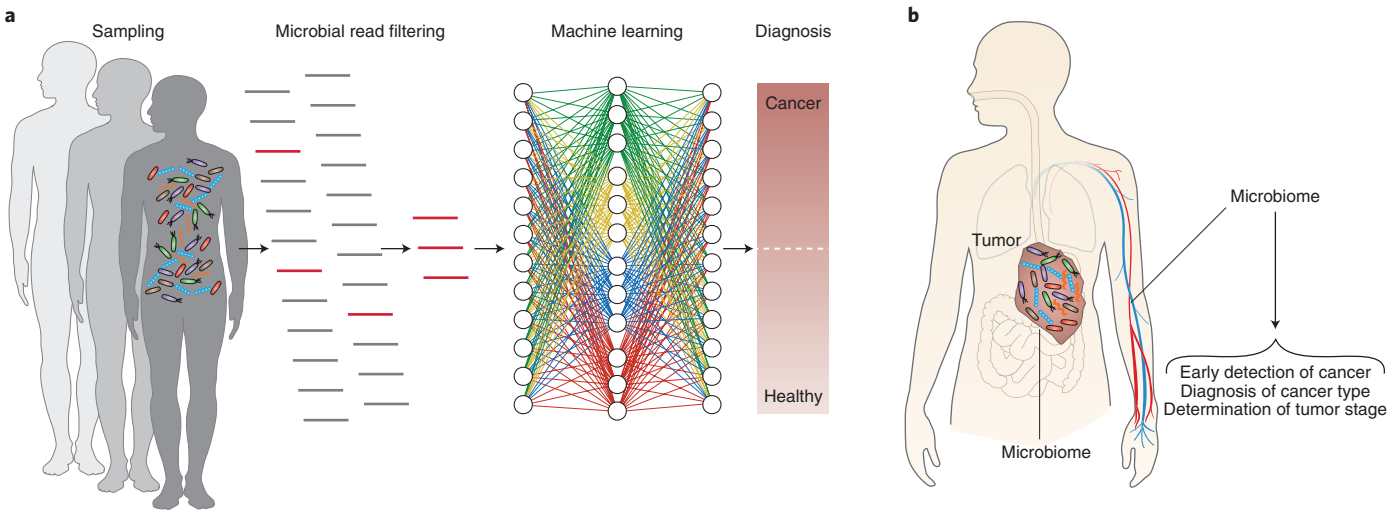

Fig. 1 | Microbial DNA signatures from tissue and blood to classify cancer. a, Tissue or blood samples from cohorts of patients with cancer and healthy controls are sequenced. The authors classified microbial DNA signatures and used them to build up machine-learning models that can be used for cancer diagnosis. **b**, Patients with cancer carry specific microbial DNA signatures that could potentially be used for early detection as well as determination of tumor type and stage in the clinic.

microbial DNA in sequencing data from patients with cancer (Fig. 1). Notably, the authors shared their algorithms and performance measures in an open-access web-based data browser that enables others to further build on their important findings (http://cancermicrobiome.ucsd.edu/CancerMicrobiome_DataBrowser/). Established associations provided by the authors between microbial signatures and clinical metadata will be likely prospectively probed in future mechanistic studies aimed at harnessing their findings for human cancer diagnostics and treatment. Additionally, the approach from Knight and colleagues may lead to the elucidation of diverse disease-associated blood microbiome signatures beyond cancer, which will pave the way for personalized, rational interventions into a complex array of human pathologies.

From the clinic: Tim F. Greten and Firouzeh Korangy

Although the comprehensive characterization of the molecular underpinnings of cancer has focused primarily on alterations in the human genome, recent reports have also identified an association of specific cancer types with specific microbiome landscapes[5,10]; however, a detailed understanding of the extent of these interspecies associations and their impact on cancer diagnosis, prognosis and treatment is still lacking. Based on the knowledge that bacteria and viruses are associated with certain types of cancer, Knight and colleagues[4] took the initiative to re-examine available whole-genome and RNA sequencing data from more than 18,000 tumor samples across 33 cancer types obtained from over 10,000 patients from TCGA. After stringent filtering and classification, the authors used the microbial DNA sequences found in these samples to develop an algorithm that allowed them to distinguish tumor and normal tissue for 15 different cancers and also to classify specific cancer types. Tissue-based microbiome models were used to discriminate between stage I and IV cancers, which worked well for colon, gastric and renal cancers but displayed a more limited performance for other cancer types. More in-depth analysis of samples derived from patients with colon, gastric, cervical, head and neck squamous cell and primary liver cancers established the biological relevance of these microorganism genetic profiles, with *Fusobacterium* spp. present in colon cancer, *Alphapapillomavirus* genus in cervical cancer, head and neck squamous cell carcinoma, hepatitis B (HBV) in HBV-infected hepatocellular carcinoma and Epstein-Barr virus (EBV) in patients with EBV-infected primary gastric cancer.

The authors moved one step further and used sequencing data from TCGA blood samples to detect microbial DNA (mbDNA) signatures and to test whether their algorithm could predict the presence of different types of cancer. Remarkably, application of this TCGA-trained machine-learning classifier on the blood-sample-derived mbDNA signatures correlated these with presence of cancer. The authors also benchmarked their findings to existing circulating tumor (ct) DNA assays and found that mbDNA could distinguish between stages Ia and IIc cancers and tumors without detectable genomic alterations, thereby potentially providing a novel tool where conventional ctDNA approaches fall short. Finally, using metagenomics sequencing of cell-free DNA extracted from plasma samples from an independent clinical cohort confirmed that this assay could discriminate between healthy individuals and patients with cancer, but also between patients with prostate cancer and lung cancer. In summary, by performing an in-depth characterization of tumor and circulating mbDNA signatures, the authors provide an innovative approach to separate patients with cancer from healthy individuals, discriminate specific cancer types and, in some cases, to identify patients with early stage versus those with more advanced disease.

This study complements a number of past reports highlighting the significance of the microbiome in cancer[11]. Data from the Human Microbiome project have demonstrated that microbial signatures display spatial and temporal variation. This diversity remains largely unexplained, although diet, environment, host genetics and early microbial exposure are all contributing factors to this heterogeneity[12]. More recently, the presence of bacteria in different primary cancers, many of them arising in the gastrointestinal tract, has been reported to influence cancer outcomes[5,10]. These findings are not limited to primary tumors, but interestingly the same bacteria

can also be found in metastasis from the same patient demonstrating microbiome stability between paired primary and metastatic tumors[5]. Although our knowledge on the biological function of the microbiome in the context of cancers remains rather limited, we do know that it can influence patients' outcome[10] as well as response to conventional chemotherapy and immunotherapy[13].

The gold standard (and only) technique to diagnose cancer depends on a tissue biopsy that is then microscopically examined by a pathologist. A genetic analysis of the patients' cancer is routinely added to find genetic alterations, which may help guide decisions regarding targeted treatment options. Liquid biopsies are a promising approach to study circulating tumor cells or perform analysis of ctDNA but are mainly used to monitor patients with an already established diagnosis, during treatment or in the adjuvant setting after surgery. Finally, the Immunoscore developed by Galon and Fridman can be used to determine the risk of relapse in early stage colon cancer patients by measuring the host immune response at the tumor site[14].

To understand the potential clinical implications of using microbial DNA for the benefit of patients with cancer, it is important to place the findings of Knight and colleagues in the context of these existing approaches for cancer diagnosis and clinical follow-up. Viewed through that lens, the authors made a number of very intriguing observations: first, they were able to identify patients with early stage cancers and without detectable genomic variations, as evidenced by ctDNA analysis; second, in some cases, they were able to correctly assign patients to specific cancer stages, outperforming alternative liquid biopsy methods. These findings suggest that the mbDNA assay may be able to identify patients with early (and still curable) disease and could be used as a rapid cancer screening tool.

Despite the clear clinical potential of these findings, some outstanding questions remain. For instance, whether the microbial signatures can predict outcome as well as the risk to develop cancer, remains unclear. Prospective studies will be needed to provide definitive proof of the utility of this potentially powerful approach. Another interesting question is whether the microbial signatures would actually change upon treatment, in which case they could represent a predictive biomarker for patient outcome and monitoring. Clearly, further testing and validation of this approach is warranted, and thus it will still take some time until a test might be ready for regulatory approval and use in clinical settings; the exciting data of Knight and colleagues call for more follow-up studies in larger cohorts and patients with different types of cancers. ❐

Lorenz Adlung[1], Eran Elinav[1,2]✉, Tim F. Greten[3]✉ and Firouzeh Korangy[3]

[1]Immunology Department, Weizmann Institute of Science, Rehovot, Israel. [2]Cancer-Microbiome Division Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany. [3]Thoracic and Gastrointestinal Oncology Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA.

✉*e-mail: eran.elinav@weizmann.ac.il; tim.greten@nih.gov*

Published online: 20 April 2020
https://doi.org/10.1038/s43018-020-0059-x

References

1. Viaud, S. et al. *Science* **342**, 971–976 (2013).
2. Sivan, A. et al. *Science* **350**, 1084–1089 (2015).
3. Gopalakrishnan, V. et al. *Science* **359**, 97–103 (2018).
4. Poore, G. D. et al. *Nature* **579**, 567–574 (2020).
5. Bullman, S. et al. *Science* **358**, 1443–1448 (2017).
6. Eisenhofer, R. et al. *Trends Microbiol.* **27**, 105–117 (2019).
7. Geller, L. T. et al. *Science* **357**, 1156–1160 (2017).
8. Zeevi, D. et al. *Cell* **163**, 1079–1094 (2015).
9. Nemet, I. et al. *Cell* **180**, 862–877.e22 (2020).
10. Riquelme, E. et al. *Cell* **178**, 795–806.12 (2019).
11. Ma, C. et al. *Science* **360**, eaan5931 (2018).
12. Human Microbiome Project Consortium. *Nature* **486**, 207–214 (2012).
13. Roy, S. & Trinchieri, G. *Nat. Rev. Cancer* **17**, 271–285 (2017).
14. Pagès, F. et al. *Lancet* **391**, 2128–2139 (2018).

Acknowledgements

L.A. received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No. 842203. E.E. is the incumbent of the Sir Marc and Lady Tania Feldmann Professorial Chair, a senior fellow at the Canadian Institute of Advanced Research (CIFAR) and an international scholar at the Bill & Melinda Gates Foundation and the Howard Hughes Medical Institute. T.F.G. is supported by the Intramural Research Program of the NIH, NCI (ZIA BC 011344).

Competing interests

E.E. is a scientific consultant for DayTwo and BiomX.