

Review



Cite this article: Zukowski A, Rao S, Ramachandran S. 2020 Phenotypes from cell-free DNA. *Open Biol.* **10**: 200119. <http://dx.doi.org/10.1098/rsob.200119>

Received: 5 May 2020

Accepted: 11 August 2020

Subject Area:

bioinformatics/genomics/molecular biology/
systems biology

Keywords:

cell-free DNA, chromatin dynamics,
subnucleosomes, cancer biomarker,
DNA methylation

Author for correspondence:

Srinivas Ramachandran

e-mail: srinivas.ramachandran@cuanschut.edu

[†]These authors contributed equally to this study.

Phenotypes from cell-free DNA

Alexis Zukowski[†], Satyanarayan Rao[†] and Srinivas Ramachandran

RNA Bioscience Initiative, and Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Mail Stop: 8101, 12801 East 17th Avenue L18–9102, Aurora, CO 80045, USA

SR, 0000-0003-2929-1377

Cell-free DNA (cfDNA) has the potential to enable non-invasive detection of disease states and progression. Beyond its sequence, cfDNA also represents the nucleosomal landscape of cell(s)-of-origin and captures the dynamics of the epigenome. In this review, we highlight the emergence of cfDNA epigenomic methods that assess disease beyond the scope of mutant tumour genotyping. Detection of tumour mutations is the gold standard for sequencing methods in clinical oncology. However, limitations inherent to mutation targeting in cfDNA, and the possibilities of uncovering molecular mechanisms underlying disease, have made epigenomics of cfDNA an exciting alternative. We discuss the epigenomic information revealed by cfDNA, and how epigenomic methods exploit cfDNA to detect and characterize cancer. Future applications of cfDNA epigenomic methods to act complementarily and orthogonally to current clinical practices has the potential to transform cancer management and improve cancer patient outcomes.

1. Introduction

Blood is a minimally invasive source for tracking an individual's health status. Most known biomolecules found in blood, such as proteins, DNA, RNA, lipids, and metabolites inform us of some aspect of body function. However, using blood to diagnose and track cancer is still a major challenge. Minimally invasive diagnostics for cancer can greatly reduce pain and suffering of patients, and if cheaper than current methods, could be performed more often in the course of treatment to monitor disease state and inform clinical care. Genomic characterization of cancer either from biopsies or plasma cell-free DNA (cfDNA) has the added potential of providing personalized treatment options.

cfDNA is DNA bound to mononucleosomes and is found circulating extracellularly in the blood. cfDNA was first identified in 1948 [1] in the blood of patients; however, the hypothesis that cfDNA acts as a reflection of disease state arose from observations in 1977 [2]. This relationship between cfDNA and disease has been a subject of study ever since. There are multiple possible pathways for the release of DNA fragments from cells, including apoptosis, necrosis, and exosome secretion [3–6]. Processes that increase the release of cfDNA include disease, inflammation, tissue injury, and exercise [7,8]. In healthy individuals, haematopoietic maturation is a major contributor to the normal cfDNA pool. Lui *et al.* elegantly demonstrated that lymphoid/myeloid tissues are the major contributors to cfDNA by identifying Y-chromosome sequences in plasma of female recipients of bone marrow transplantations from male donors [9]. Multiple studies since then further confirmed that the lymphoid/myeloid tissues mainly contribute to the normal cfDNA pool [10–13]. Tumours, when present, also contribute to cfDNA. Thus, cfDNA is a molecular barcode of cells undergoing turnover and is an attractive target for clinical diagnostics and real-time monitoring of many cancers.

Provided that there are ways to distinguish cfDNA originating at the disease site from cfDNA produced during normal turnover, cfDNA could be used for diagnosis. The major focus of using cfDNA for cancer diagnosis has been the identification of a limited disease-specific panel of mutations. However, mutation

detection has a significant limitation: that clonal haematopoiesis contributes to a significant fraction of mutations in cfDNA, including mutations in prominent cancer-associated genes like TP53 and DNMT3A [14]. Moreover, these mutations in cfDNA from the early stages of disease can be indistinguishable from mutations that arise at appreciable rates in healthy tissues [15,16]. This major obstacle has led to the exploration of other orthogonal information in cfDNA that could identify its tissue-of-origin and, in turn, disease states. Epigenomes reflect cellular identity and phenotype, and our ability to connect epigenomes to cellular identity could be used to infer disease phenotypes from cfDNA. Significantly, epigenome-based cfDNA approaches would be orthogonal and complementary to mutation-based cfDNA assays. In this review, we discuss epigenome features that can be characterized in cfDNA, and which provide information on the cfDNA tissue-of-origin.

2. Cell-free DNA contains a map of the chromatin state in the tissue-of-origin

The periodicity of cytoplasmic DNA released during red blood cell maturation in mouse fetal liver led to the hypothesis that there was a regular arrangement of protein protections on the genome [17,18]. This work preceded both the field of apoptosis and the biochemical characterization of the nucleosome and provided the first hint that cfDNA could be a map of chromatin structure. Protein-bound DNA lasts longer than naked DNA in serum [19], where nucleases are abundant, which suggests that cfDNA is probably double-stranded and protein-bound to inhibit degradation by endogenous nucleases in plasma. This is supported by the ability to detect nucleosomes in plasma by sandwich enzyme-linked immunosorbent assays (ELISA) and by chromatin immunoprecipitation (ChIP) [20–24].

Apoptosis is one of the processes that could lead to genome fragmentation into small protein-DNA complexes that are found in circulation. Apoptosis involves the upregulation of nucleases that attack the cell's genome. Characteristic of this 'attack' is DNA fragmentation that results in a laddering of repeat species less than 5 kb [5,25]. The repeating unit in laddering seen in apoptosis corresponds to approximately 150 bp, which parallels the length of DNA wrapped around a nucleosome. Therefore, plasma cfDNA, which is typically in the form of a mononucleosome, informs us about the chromatin structure of cells undergoing turnover. In other words, the 'epigenome' of the tissue-of-origin can be measured from cfDNA (figure 1).

3. Structural epigenomics

It has long been known that chromatin structure in cells can be inferred from digestion patterns of nucleases that are sensitive to protein-DNA contacts [26]. Micrococcal nuclease (MNase) has been used to study chromatin structure for decades [27]. Understanding chromatin structure by sequencing fragments protected from MNase digestion has an unexpected parallel in cfDNA. The information derived from fragment length and genome location can inform us of biochemical activity at a genomic locus for MNase and cfDNA. For cfDNA, this can be used to infer transcriptional activity and tissue-of-origin. As we will demonstrate below, the length and genomic location

of fragments protected from a nuclease can tell us the locus-specific distribution of nucleosomes, transcription factors (TF), and nucleosomal intermediates formed during transcription (figure 1). The ability to identify locus-specific structures of protein-DNA complexes from sequencing data, termed 'structural epigenomics', is a general strategy applicable to a wide variety of datasets, including cfDNA sequencing datasets [11].

4. Nucleosome positions reflect genome function

Eukaryotic genomic DNA is packaged with nucleosomes like 'beads on a string' and consecutive nucleosomes are separated by short linker DNA [28,29]. MNase preferentially degrades linker DNA and is inhibited when it encounters a protein-DNA contact [27,30]. A nucleosome protects 147 bp of DNA, a chromosome (nucleosome with a linker histone bound) protects 167 bp of DNA, and TFs protect less than 50 bp of DNA. MNase has traditionally been used to map positions of whole nucleosomes in intact nuclei by purifying approximately 147 bp DNA after MNase digestion and subjecting this DNA to massively parallel short-read sequencing [31]. Nucleosome positioning impacts all biochemical processes that occur on the genome and consequently, knowing nucleosome positions enables us to predict biochemical activities that occurred in the cells that resulted in these protections [32]. A striking example of a distinct nucleosome organization is found in active genes. There is a depletion of nucleosomes at active promoters and nucleosomes are well-ordered upstream and downstream of active promoters. cfDNA contains information on nucleosome positioning and high-quality nucleosome maps were obtained from deep, whole-genome cfDNA sequencing from the plasma of healthy donors [12,33–35].

Nucleosome density inferred from cfDNA of healthy individuals had the same features as that of lymphoid cell lines: genes highly expressed in lymphoid cell lines featured nucleosome depletion at promoters and ordered nucleosome positions upstream and downstream of the promoter in cfDNA (figure 1). By contrast, genes that were not expressed in these cell lines showed significant nucleosome density over the promoters and lack of ordering over gene bodies in cfDNA [12,35]. These observations strongly validated the lymphoid/myeloid origin of cfDNA in healthy humans.

Quantitative scores were developed based on these striking observations to connect nucleosome profiles to gene activity in the cfDNA tissue-of-origin. Snyder *et al.* showed that stronger periodicity of nucleosomes in the gene body (the region between the transcription start site (TSS) and TSS+5000 bp) correlated with higher gene expression in lymphoid/myeloid tissues for healthy donors [12]. However, in cfDNA from donors with cancer, the correlation of nucleosome periodicity at gene bodies with gene expression of lymphoid/myeloid tissues was much weaker, suggesting other cell types contributing to the cfDNA pool. This was confirmed by the fact that the correlation of periodicity with gene expression of other cell types increased for donors with cancer [12]. Thus, nucleosome periodicity could inform the gene expression of the cfDNA tissue(s)-of-origin.

Utz *et al.* noted that overall nucleosome density was lower in 2 kb upstream and downstream of the TSS (which they termed '2 K-TSS') and lowest at the promoter itself (usually

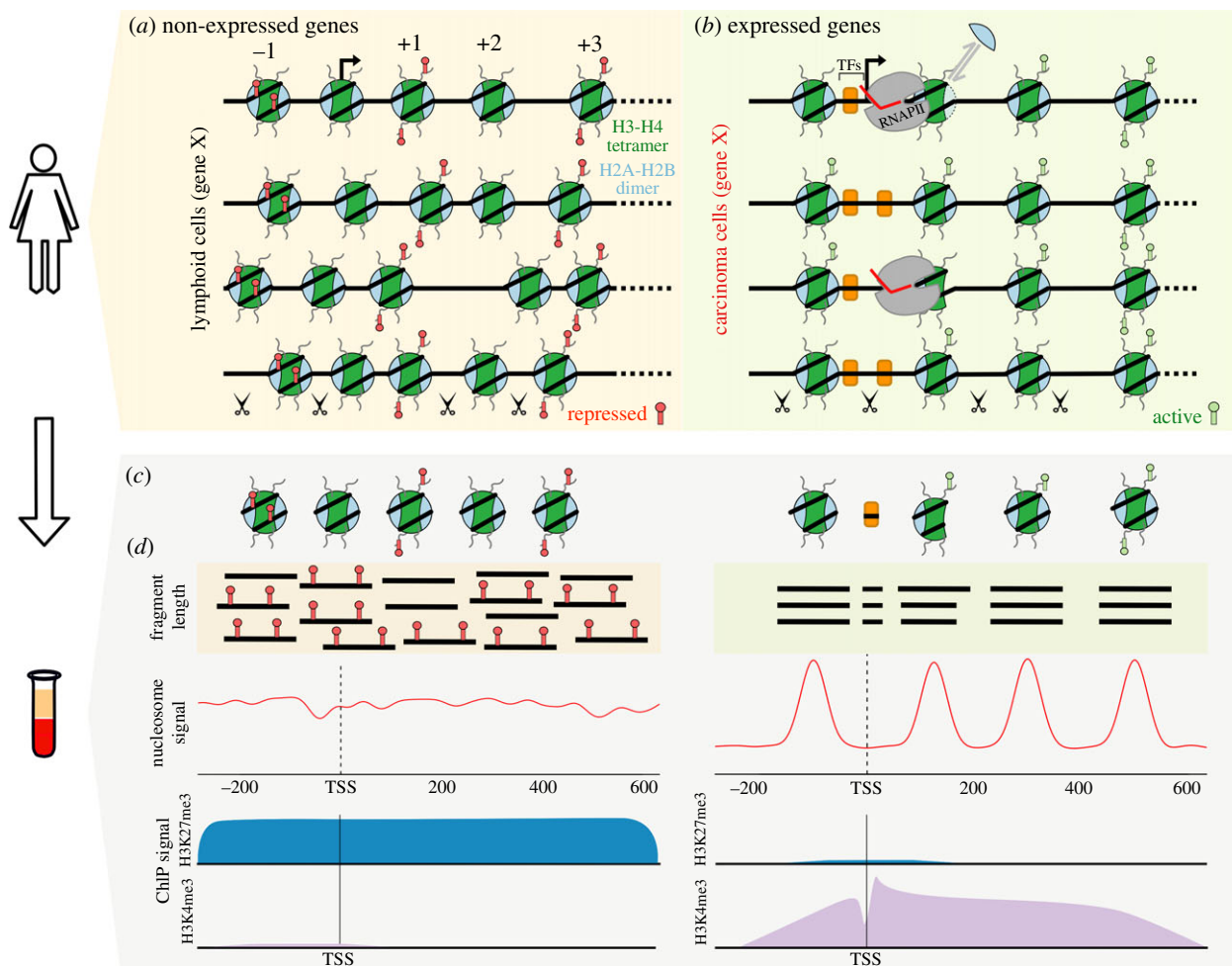


Figure 1. Cell-free DNA reflects the structural epigenomic information of the cell-of-origin. Schematic of the nucleosomal landscape differences of gene X when it is not expressed (left) and expressed (right) in different cell-types. (a) A non-expressed gene features promoter-proximal DNA methylation (red flags), methylation of nucleosomal histone tails (red flags, H3K27me3), nucleosome occlusion of the promoter (arrow), a lack of transcription factors (TFs) upstream of the promoter, and the absence of RNA polymerase II (RNAPII). (b) At expressed genes, nucleosomes are well-positioned, modifications like H3K4me3 (green flags) are present, RNAPII occupies the promoter, and TFs are bound upstream. At the +1 nucleosome during transcriptional elongation, RNAPII transiently breaks DNA-histone contacts allowing H2A-H2B dimers to exchange (light blue crescent). Nucleases, shown as scissors, are ubiquitously present and preferentially cleave accessible DNA. At gene X, when protein protections of DNA change during chromatin remodelling events, such as transcription, nuclease activity captures the different DNA length protections (see fragment lengths). (c) During cell-turnover, DNA-protein complexes are released into circulation. (d) Circulating DNA-protein complexes in plasma preserve the epigenomic features of the transcriptional status of the cell-of-origin. These features include fragment lengths (protein-DNA protections), DNA methylation, nucleosome positioning profiles, and nucleosome post-translational modifications. Note the preservation of these transcriptional and non-transcriptional DNA-protein species from in the body to plasma.

termed the ‘nucleosome depleted region’, NDR) in expressed genes in both MNase-seq of a lymphoblastoid cell line and the cfDNA sequencing data from healthy donors [35]. They developed a model that used the nucleosome occupancy at 2 K-TSS and NDR to predict gene expression. In a person with cancer, tumour-contributed cfDNA would have a depletion of nucleosomes at active genes, but this depletion may be masked by high nucleosome occupancy in cfDNA from lymphoid/myeloid tissues. Ulz *et al.* circumvented this problem by focusing on regions in the genome that featured copy-number gains in cfDNA of donors with cancer. These genomic regions would have a higher representation in tumour cfDNA if the tumour was the source of these copy number gains. Their simulations suggested that at least 75% of cfDNA at a given TSS must be released from tumour cells to infer expression status using their method. In the regions of copy number gain, they demonstrated significant changes in nucleosome occupancy at 2 K-TSS and NDR that correlated with increased expression of these genes in the tumour as

assessed by RNA-seq of matched tumour biopsy samples. Thus, promoter and gene-body depletion of cfDNA fragments are indicative of gene activity in the tissue-of-origin of cfDNA.

5. Subnucleosomes in cell-free DNA inform expression state of genes

Transcription also results in the formation of subnucleosomes proximal to the promoter and identification of subnucleosomes could serve as a proxy for measuring gene activity. In eukaryotic cells, transcription occurs on a chromatin template. Because RNA polymerase II (RNAPII) needs to unwind the DNA strands to make RNA, every protein-DNA contact in its path needs to be disrupted. *In vitro*, nucleosomes present a substantial barrier to transcription elongation [36,37], resulting in specific stalls by RNAPII at sites of strong histone-DNA contacts. In cells, we can map the positions where RNAPII stalls with base pair-resolution

maps of the 3' ends of nascent transcripts [38–41]. These maps have shown that the first nucleosome downstream of TSS (+1 nucleosome) presents a strong barrier to RNAPII in cells.

The effect of RNAPII stalling on the +1 nucleosome could be understood by mapping intermediate nucleosome states during transcription elongation. With sequencing library protocols that capture all fragment lengths combined with paired-end sequencing, the full spectrum of fragments generated by MNase treatment of nuclei can be uncovered [11,42–44]. Fragments between 50 and 147 bp are too short to be protected by a whole nucleosome, but too long to be TF footprints. These intermediate-length protections represent a discrete loss of contacts asymmetrically from either side of the nucleosome as it unwraps during transcription and remodeling (figure 1). Correlating a base-pair resolution map of RNAPII with the high-resolution distribution of nucleosomal intermediates at the +1 nucleosome revealed that nucleosome unwrapping occurs in a stepwise manner—first, the contacts to the H2A-H2B dimer proximal to the promoter are lost. Then as RNAPII elongates through the nucleosome, contacts to the H2A-H2B dimer distal to the promoter are lost [11]. Cryo-electron microscopy of unwrapped nucleosomes and of RNAPII transcribing nucleosome templates yielded structures that matched the *in vivo* structures inferred from paired-end sequencing [45–47]. These observations highlight the ability of genomic sequencing of short DNA fragments to detect transcription-dependent substructures of nucleosomes in cells.

cfDNA is highly nicked and nicked DNA is lost in standard double-stranded library preparation protocols. Inspired by methods used in Neanderthal genome sequencing projects, Snyder *et al.* used a single-stranded library protocol (SSP) that captures both nicked and non-nicked fragments [12,48]. Surprisingly, they found an abundance of fragments less than nucleosomal size, ranging from approximately 40 bp onwards. These short fragments resemble subnucleosomes observed in MNase sequencing data from *Drosophila* cells [11]. When the subnucleosomal cfDNA fragments were mapped at +1 nucleosome positions of genes expressed in lymphoid/myeloid tissues, the same asymmetric unwrapping intermediates that were seen in *Drosophila* cells could also be observed in the cfDNA data [11]. Thus, it seems that the long residence time of RNAPII near the +1 nucleosome results in long-lived nucleosomal intermediate states that are mappable both by MNase, and by the endogenous nucleases that give rise to cfDNA.

In *Drosophila* cells, it was found that the amount of the subnucleosomal particles relative to nucleosomal particles at the +1 nucleosome of a gene correlated with the extent to which that gene was transcribed [11]. Therefore, it was hypothesized that the amount of these intermediates relative to whole nucleosomes (subnucleosome enrichment) at the +1 nucleosome of genes in cfDNA would correlate with the composite expression profile of cells giving rise to cfDNA. Indeed, in a healthy donor, cfDNA subnucleosome enrichment correlated much better with expression profiles of lymphoid/myeloid tissue types compared to other tissue types. However, in donors with cancer, the subnucleosome enrichment corresponding to lymphoid/myeloid tissues weakened substantially, enabling robust differentiation of healthy and cancer plasma cfDNA [11]. Thus, subnucleosomes in cfDNA represent relics of transcription in cfDNA tissues-of-origin that can inform us about the transcriptional programmes active at disease sites.

6. Chromatin structure correlations across megabases

The observation that nucleosome dynamics are reflected by shorter protections in cfDNA can be extended to much larger length scales than single nucleosomes, which allows maximal use of low depth cfDNA sequencing data. Using a machine learning model, Cristiano *et al.* were able to observe similar ‘fragmentation profiles’ at megabase scales between cfDNA from healthy donors and the DNA released by MNase treatment of healthy lymphocytes [49]. Fragmentation profile is a measure similar to subnucleosome enrichment: a ratio of small fragments (100–150 bp) to large cfDNA fragments (151–220 bp). A significant difference in fragmentation profiles was observed between cfDNA from healthy donors and donors with cancer, enabling detection of cancer with low-depth whole-genome cfDNA sequencing. This method works presumably because fragmentation profiles at large length scales differentiate active domains of the chromosome from inactive domains. A tumour is expected to have significantly different active chromatin domains at the megabase scale compared to lymphoid/myeloid tissues.

A similar hypothesis is that the strength of correlation of cfDNA length profiles between genomic loci is proportional to the correlation of contacts made by the loci to the rest of the chromosome [50,51]. The correlation matrix of contact probability for a cell type can be obtained by Hi-C [52]. The Hi-C method involves crosslinking cells, fragmenting the genome into large pieces with the crosslinks intact, and then ligating the cross-linked fragments. Sequencing of the ligated contacts reveals chromatin contacts genome-wide. Regions of the genome with similar activity share similar contact profiles across the chromosome. Liu *et al.* showed that the strength of correlation of healthy cfDNA length profiles between different genomic regions is proportional to the strength of correlation of Hi-C contacts between different genomic regions as measured in lymphoblastoid cells. Furthermore, this correlation of cfDNA length profiles between different regions of a chromosome could be modelled as a combination of Hi-C maps of different tissue types to uncover tissue contributions to cfDNA in tumour samples with high (greater than 30%) tumour fractions [50]. Thus, cfDNA profiles could be used to infer chromosome contacts in the tissue(s)-of-origin.

7. Distinct cell-free DNA patterns at transcription factor binding sites

TF binding sites at promoters, enhancers, and insulators show enrichment of short protections (less than 50 bp) and depletion of nucleosomes in MNase-seq experiments [11,42,43,53–55]. Binding of many TFs also results in the ordering of nucleosomes around the TF binding sites [56]. Thus, protected TF binding sites represent a discrete class of genomic loci that show distinctive chromatin structural profiles in nuclease-protection assays. The enrichment of short fragments in cfDNA using SSP enabled Snyder *et al.* to observe similar enrichment of short protections and ordered nucleosome arrays at TF binding sites in cfDNA datasets from healthy plasma. They also observed an

enrichment of short fragments in cfDNA at promoters that was proportional to expression levels of a lymphoid cell line, demonstrating the ability to possibly identify transcriptional complexes at a promoter from cfDNA [12].

Ulz *et al.* found that mononucleosome depletion observed in cfDNA at known TF-binding sites (TFBSs) is a good proxy for inferring TF binding [57]. This enables inference of TF binding from cfDNA sequenced using standard double stranded protocols that do not enrich for short fragments. They selected a set of TFs which are lineage-specific, for example, Androgen Receptor (prostate), Even-Skipped Homeobox 2 (colon), Forkhead box A1 (breast), and for each of them, profiled 1000 binding sites that were concordant across tissue samples [58]. Using plasma from patients with prostate, breast, and colon cancer, they showed that nucleosome depletion levels at binding sites of tumour-specific TFs are significantly higher compared to healthy cohorts and the reverse was true at binding sites for haematopoietic lineage-specific TFs such as LYL1, and EVI1. In addition, they also showed that nucleosome depletion levels at specific TFBSs were predictive of tumour subtypes. For instance, they found a significant reduction in nucleosome depletion at Androgen Receptor-binding sites when comparing samples from before and after a patient's prostate adenocarcinoma became androgen-independent. Thus, TF binding can be inferred from cfDNA either directly through short fragment protections when using SSP or indirectly via nucleosome depletion at TFBSs when using standard double-stranded library protocols and tissue-specific TFBSs enable identification of tissues that contribute to cfDNA.

8. Cell-free DNA fragment length

As we have discussed above, cfDNA length reflects the structure of chromatin in cells of origin and can be explicitly modelled as such. Beyond trying to understand cfDNA profiles in terms of chromatin structure, there have been several useful observations about the use of cfDNA length in biomarker development, for which we can only speculate the underlying molecular details. At its most general, the size of cfDNA fragments is used to delineate between baseline homeostasis and other biological phenomena that give rise to cfDNA. cfDNA has already been directly applied in the clinic for pre-natal testing. In plasma, fetal cfDNA is shorter than maternal cfDNA [13,59,60], which enables detection of aneuploidy or recessive disorders in the fetus non-invasively. Urine cfDNA features short periodic fragments that resemble subnucleosomes [51]. In individuals with cancer, circulating tumour DNA resides in the shorter cfDNA fraction [61–63]. The frequency of shorter, non-tumour cfDNAs is typically low and thus size selection for shorter fragments could dramatically increase the sensitivity of detection for mutant tumour cfDNA. Mutation targeting and identification of copy number variation and/or single copy number alterations is achieved with a higher sensitivity if the shorter cfDNA fraction is profiled [61,62]. This is a consequence of mutant tumour cfDNA being diluted less with non-tumour cfDNA in the shorter fraction. In the light of these observations, one prediction is that shorter cfDNA is a product of a specific mechanism that occurs at a higher frequency in fetal cells, cells that give rise to cfDNA in urine, and cancer cells relative to normal lymphoid/myeloid cell turnover.

9. Cell-free DNA chromatin immunoprecipitation uses cell type-specific genome-wide patterns of histone modifications to identify tissues of origin

Several studies in the early 2000s used ELISA to capture and calculate the concentration of nucleosomes in plasma [20,22–24,64,65]. Specifically, nucleosomes were captured in an antibody 'sandwich', which used an antibody for a histone and an antibody for double-stranded DNA. This approach was an alternative means to assess cell death using plasma from cancer patients. In a notable study, cancer patients exhibited a statistically significant increase in nucleosome concentration compared to healthy individuals [20]. These findings suggest that nucleosome concentration could act as a biomarker for a disease state and/or discriminate between individuals with disease and those who are healthy. However, this application lacks specific genomic information to identify the tissue-of-origin and specify the disease state.

Beyond differences in overall nucleosome concentrations among physiological states, an intriguing question is if post-translational modifications (PTMs) of nucleosomal histones in plasma can be used to differentiate healthy individuals and those with disease (figure 1). PTMs of nucleosomal histones and DNA differ in euchromatic regions versus heterochromatic regions [66,67]. PTMs range from small chemical groups to larger proteins, such as ubiquitin, that are added to specific amino acids of proteins. For example, tri-methylation of histone 3 lysine 27 (H3K27me3) or H3K9me3 are found at heterochromatic regions where gene activity is repressed or 'silent'. By contrast, H3K4me3 or H3K36me3 are associated with euchromatic regions where genes are active [68,69].

The link between cancer and chromatin regulation is well documented [70]. The majority of oncogenic mutations occur in genes of chromatin modifiers [71–73], which include a number of tumour-suppressor genes, such as SIRT1 [74]. Genomic instability is a hallmark of cancer and a consequence of alterations in heterochromatin that disrupt silencing [75,76]. Deligezer *et al.* analysed histone PTMs on circulating nucleosomes in plasma [22]. In light of the evidence for deregulation of repressive PTMs associated with gene silencing in cancer, histone methylation, specifically H3K9me1, was first assessed on circulating nucleosomes from myeloma patient plasma samples. Follow-up papers identified a reduction in repressive chromatin marks (H3K9me3, H4K20me3 and H3K27me3) in patients with colorectal cancer and metastatic prostate cancer [22,23]. However, these studies did not use high-throughput sequencing technology that would provide information on the genomic locations of histone PTMs; an aspect that is important to consider when delineating between different physiological states (i.e. disease and non-disease states) that produce changes in PTM levels in circulating nucleosomes.

A recent study sought to assess circulating post-translationally modified nucleosomes in a variety of physiological states using a combination of ChIP and next-generation sequencing [21]. The authors created a workflow to isolate modified circulating nucleosomes from plasma and performed high-throughput sequencing of the associated DNA (cfChIP-seq). Modified nucleosomes containing PTMs associated with active genes or *cis*-regulatory elements (i.e. enhancers) were

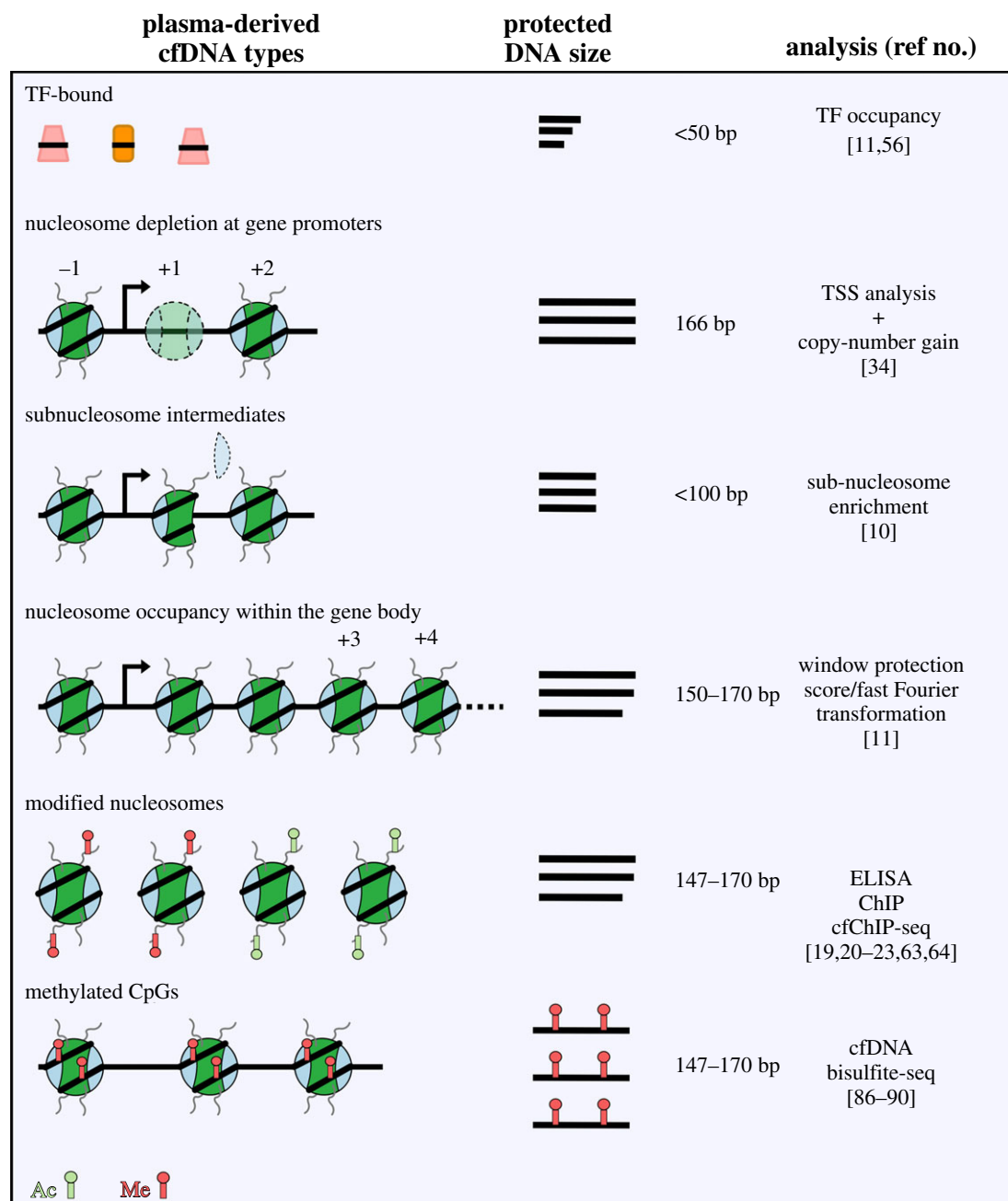


Figure 2. Clinical applications for cancer patients using epigenomic features of cell-free DNA.

specifically targeted: H3K4me1/2/3 and H3K36me3. H3K4me3 cfChIP-seq signal was used as a proxy for active genes and recapitulated earlier findings that cell-type specific gene activation programmes belonged to mainly blood lineages in a healthy individual. Furthermore, the cfChIP-seq signals reflected previously identified ChIP-seq signals for PTMs in these lineages.

This analysis was extended to profile patients with a variety of physiological insults, including surgery and cancer. cfChIP-seq detected differences in gene expression in individuals with cancer compared to healthy individuals. In response to surgical interventions, changes in tissue-specific cfChIP-signatures were detectable in the blood of patients supporting tissue-of-origin specificity. Combinations of PTMs were also assessed for complementary information to further profile tissue-type specific gene activity from non-coding regions. The authors observed changes in H3K4me2 (found at putative enhancer elements) and H3K36me3 (found within the body of transcribed genes) that correlated with differential

gene activity unique to cancer tumours. Overall, the ability to identify the tissue-of-origin based on enrichment of circulating nucleosome PTMs associated with transcription at specific regions of the genome is exciting. cfChIP-sequencing provides an additional layer of information about an individual's underlying physiological state derived from the blood.

10. Cell-free DNA methylation patterns identify tissues of origin

Methylation of cytosines in the CpG dinucleotide context is associated with cellular identity [77] and can be identified on cfDNA [78,79]. Deamination of cytosine to uracil using sodium bisulfite or using an antibody specific to CpG methylation to pulldown methylated DNA are the two main methods used to map CpG methylation genome wide. Often promoter hypermethylation of a gene is associated with silencing [79]. DNA methylation profiles are stable and cell-type specific,

and differentially methylated regions (DMRs) have been extensively used in quantification of subpopulations in DNA extracted from heterogeneous cell populations [80,81]. Accomando *et al.* showed that cell-types in leucocytes can be quantified using methylation status at as few as 20 CpG loci [82]. Many genomic locations exhibit highly cell-type specific DNA methylation, which has been instrumental in developing computational methods to delineate the contribution of tissues to cfDNA. CpG islands, a region 300–3000 bp in length and rich in G + C content, are usually hypomethylated, but aberrant methylation of these regions is associated with diseases like cancer [83]. Some CpG islands share high methylation levels across tumour types, while others are tumour specific [84]. Promoter hypermethylation of tumour suppressor genes is especially a striking feature of tumours particularly during carcinogenesis [85,86]. Targeted amplification of candidate promoter regions has been used to find that cfDNA from cancer patients exhibit hypermethylation in contrast to healthy donors [87–89]. Using methylated DNA immunoprecipitation (MeDIP-seq) optimized for low DNA amounts, Shen *et al.* detected numerous DMRs in cfDNA from individuals with cancer that are specific to tumour tissue-of-origin [90]. In particular, they showed that plasma DMRs in individuals with colorectal cancer closely matched solid tumour-derived DMRs. Thus, tissue-specific methylation patterns and hypermethylation of tumour suppressor genes can be used as signatures to detect cancer as well as identify cfDNA tissue(s)-of-origin [91].

11. Conclusion/future perspectives

The remarkable observation in 1970 that hinted at a relationship between cytoplasmic DNA and the chromatin structure of the cell-of-origin now underlies the conceptual paradigm of cfDNA epigenomics [18]. As there is an increasing awareness of the limitations associated with detecting and identifying diseases, such as cancer, solely based on mutant genotypes, the emergence of cfDNA approaches based on epigenomics provides a valuable alternative (figure 2). Chromatin remodellers, chromatin-modifying complexes and DNA methyltransferases are frequently dysregulated in cancer highlighting the relationship between chromatin regulation and oncogenesis [92–97]. Furthermore, simulations show that using hundreds of features throughout the genome (which would be the case for epigenomic features) improves the limit of detection of disease states from cfDNA compared to a few features (which would be the case when looking for tumour mutations) [49,90]. Combining epigenomic

assessments with mutation-based targeting may better guide personalized treatment of cancer patients rather than solely relying on one method.

Early detection of cancer still remains a challenge for all cfDNA-based methods. An early cancer diagnosis increases a patient's chance for curative treatment. However, tumour shedding of DNA into the blood is at a lower concentration in early stages than what is observed in advanced cancer stages. Most methods described above have been used on samples of advanced cancer patients indicating a much-needed effort to examine their use and sensitivity for early detection. However, significant in-roads are being made. A recent multi-centre, randomized study used cfDNA methylation analysis to prospectively assess the largest cohort ($n = 6689$, 4207 healthy and 2482 cancer) to date for early detection and localization of more than 50 types of cancers [91]. With 99.3% specificity, the sensitivity increased as cancer stage increased (Stage I: 18%—Stage IV: 93%), with tissue-of-origin predicting cancer in 96% of samples with an accuracy of 93%. These are promising results for the future development of a cancer screening test for the general population. With improvements over time, cfDNA epigenomic methods described in this review can be combined with current clinical practices to improve detection of cancer in patients.

Obtaining molecular signatures of cancer based on epigenomic features highlights the variety of information contained in the blood of an individual. The studies we reviewed here represent an orthogonal network of disease-specific epigenomic information that is rich for exploitation. Combined with current clinical molecular diagnostics, the timely emergence of epigenomic cfDNA methods herald a chance to revolutionize disease management, specifically cancer. One can envision the successful implementation of these approaches to help guide personalized medicine at different stages of disease beginning with diagnosis and throughout the course of treatment. Beyond biomarkers, diverse epigenomic cfDNA methodologies also provide the opportunity to directly investigate the molecular mechanisms underlying pathology in humans.

Data accessibility. This article has no additional data.

Authors' contributions. A.Z., S.Rao. and S.Ram. wrote the text. A.Z. and S.Rao. created the illustrations. A.Z. and S.Rao. contributed equally.

Competing interests. We declare we have no competing interests.

Funding. This work was supported by the RNA Bioscience Initiative, University of Colorado School of Medicine.

Acknowledgements. We thank Dr Olivia Rissland and Dr Richard Davis for critical comments on the manuscript.

References

- Mandel P, Metais P. 1948 Les acides nucléiques du plasma sanguin chez l'homme. *C R Seances Soc. Biol. Fil.* **142**, 241–243.
- Leon SA, Shapiro B, Sklaroff DM, Yaros MJ. 1977 Free DNA in the serum of cancer patients and the effect of therapy. *Cancer Res.* **37**, 646–650.
- Aucamp J, Bronkhorst AJ, Badenhorst CPS, Pretorius PJ. 2018 The diverse origins of circulating cell-free DNA in the human body: a critical re-evaluation of the literature. *Biol. Rev. Camb. Phil. Soc.* **93**, 1649–1683. (doi:10.1111/brv.12413)
- Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, Knippers R. 2001 DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. *Cancer Res.* **61**, 1659–1665.
- Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. 2001 About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. *Clin. Chim. Acta* **313**, 139–142. (doi:10.1016/S0009-8981(01)00665-9)
- Zhivotovsky B, Samali A, Orrenius S. 2001 Determination of apoptosis and necrosis. *Curr. Protoc. Toxicol. Ch.* **2**, Unit 2.2. (doi:10.1002/0471140856.tx0202s00)
- Paunel-Görgülü A *et al.* 2017 cfDNA correlates with endothelial damage after cardiac surgery with

- prolonged cardiopulmonary bypass and amplifies NETosis in an intracellular TLR9-independent manner. *Sci. Rep.* **7**, 17421. (doi:10.1038/s41598-017-17561-1)
8. Tug S, Helmig S, Deichmann ER, Schmeier-Jurchott A, Wagner E, Zimmermann T, Radsak M, Giacca M, Simon P. 2015 Exercise-induced increases in cell free DNA in human plasma originate predominantly from cells of the haematopoietic lineage. *Exerc. Immunol. Rev.* **21**, 164–173.
 9. Lui YY, Chik KW, Chiu RW, Ho CY, Lam CW, Lo YM. 2002 Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin. Chem.* **48**, 421–427. (doi:10.1093/clinchem/48.3.421)
 10. Lam WKJ *et al.* 2017 DNA of erythroid origin is present in human plasma and informs the types of anemia. *Clin. Chem.* **63**, 1614–1623. (doi:10.1373/clinchem.2017.272401)
 11. Ramachandran S, Ahmad K, Henikoff S. 2017 Transcription and remodeling produce asymmetrically unwrapped nucleosomal intermediates. *Mol. Cell.* **68**, 1038–1053. (doi:10.1016/j.molcel.2017.11.015)
 12. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. 2016 Cell-free DNA comprises an *in vivo* nucleosome footprint that informs its tissues-of-origin. *Cell* **164**, 57–68. (doi:10.1016/j.cell.2015.11.050)
 13. Sun K *et al.* 2015 Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. *Proc. Natl Acad. Sci. USA* **112**, E5503–E5512. (doi:10.1073/pnas.1508736112)
 14. Razavi P *et al.* 2019 High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. *Nat. Med.* **25**, 1928–1937. (doi:10.1038/s41591-019-0652-7)
 15. Martincorena I *et al.* 2018 Somatic mutant clones colonize the human esophagus with age. *Science* **23**, 911–917. (doi:10.1126/science.aau3879)
 16. Salk JJ *et al.* 2019 Ultra-sensitive TP53 sequencing for cancer detection reveals progressive clonal selection in normal tissue over a century of human lifespan. *Cell Rep.* **28**, 132–144. (doi:10.1016/j.celrep.2019.05.109)
 17. Henikoff S, Church GM. 2018 Simultaneous discovery of cell-free DNA and the nucleosome ladder. *Genetics* **209**, 27–29. (doi:10.1534/genetics.118.300775)
 18. Williamson R. 1970 Properties of rapidly labelled deoxyribonucleic acid fragments isolated from the cytoplasm of primary cultures of embryonic mouse liver cells. *J. Mol. Biol.* **51**, 157–168. (doi:10.1016/0022-2836(70)90277-9)
 19. Yao W, Mei C, Nan X, Hui L. 2016 Evaluation and comparison of *in vitro* degradation kinetics of DNA in serum, urine and saliva: a qualitative study. *Gene* **590**, 142–148. (doi:10.1016/j.gene.2016.06.033)
 20. Holdenrieder S, Stieber P, Bodenmuller H, Busch M, Von Pawel J, Schalhorn A, Nagel D, Seidel D. 2001 Circulating nucleosomes in serum. *Ann. N Y Acad. Sci.* **945**, 93–102. (doi:10.1111/j.1749-6632.2001.tb03869.x)
 21. Sadeh R *et al.* 2019 ChIP-seq of plasma cell-free nucleosomes identifies cell-of-origin gene expression programs. *Biorxiv.* 638643. (doi:10.1101/638643)
 22. Deligezer U, Akisik EE, Erten N, Dalay N. 2008 Sequence-specific histone methylation is detectable on circulating nucleosomes in plasma. *Clin. Chem.* **54**, 1125–1131. (doi:10.1373/clinchem.2007.101766)
 23. Deligezer U, Yaman F, Darendeliler E, Dizdar Y, Holdenrieder S, Kovancilar M, Dalay N. 2010 Post-treatment circulating plasma BMP6 mRNA and H3K27 methylation levels discriminate metastatic prostate cancer from localized disease. *Clin. Chim. Acta.* **411**, 1452–1456. (doi:10.1016/j.cca.2010.05.040)
 24. Holdenrieder S *et al.* 2001 Nucleosomes in serum of patients with benign and malignant diseases. *Int. J. Cancer* **95**, 114–120. (doi:10.1002/1097-0215(20010320)95:2<114::AID-IJC1020>3.0.CO;2-Q)
 25. Samejima K, Earnshaw WC. 2005 Trashing the genome: the role of nucleases during apoptosis. *Nat. Rev. Mole. Cell Biol.* **6**, 677–688. (doi:10.1038/nrm1715)
 26. Weintraub H, Groudine M. 1976 Chromosomal subunits in active genes have an altered conformation. *Science (New York N.Y.)* **193**, 848–856. (doi:10.1126/science.948749)
 27. Noll M. 1974 Subunit structure of chromatin. *Nature* **251**, 249–251. (doi:10.1038/251249a0)
 28. Kornberg RD. 1974 Chromatin structure: a repeating unit of histones and DNA. *Science* **184**, 868–871. (doi:10.1126/science.184.4139.868)
 29. Olins AL, Olins DE. 1974 Spheroid chromatin units (v bodies). *Science* **183**, 330–332. (doi:10.1126/science.183.4122.330)
 30. Noll M, Kornberg RD. 1977 Action of micrococcal nuclease on chromatin and the location of histone H1. *J. Mol. Biol.* **109**, 393–404. (doi:10.1016/S0022-2836(77)80019-3)
 31. Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K. 2008 Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**, 887–898. (doi:10.1016/j.cell.2008.02.022)
 32. Hughes AL, Rando OJ. 2014 Mechanisms underlying nucleosome positioning *in vivo*. *Annu. Rev. Biophys.* **43**, 41–63. (doi:10.1146/annurev-biophys-051013-023114)
 33. Ivanov M, Baranova A, Butler T, Spellman P, Mileyko V. 2015 Non-random fragmentation patterns in circulating cell-free DNA reflect epigenetic regulation. *BMC Genomics* **16**(Suppl. 13), S1. (doi:10.1186/1471-2164-16-s13-s1)
 34. Sun K *et al.* 2019 Orientation-aware plasma cell-free DNA fragmentation analysis in open chromatin regions informs tissue of origin. *Genome Res.* **29**, 418–427. (doi:10.1101/gr.242719.118)
 35. Ulz P *et al.* 2016 Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nat. Genet.* **48**, 1273–1278. (doi:10.1038/ng.3648)
 36. Kireeva ML, Walter W, Tchernajenko V, Bondarenko V, Kashlev M, Studitsky VM. 2002 Nucleosome remodeling induced by RNA polymerase II. *Mol. Cell* **9**, 541–552. (doi:10.1016/s1097-2765(02)00472-0)
 37. Kulaeva OI, Hsieh FK, Studitsky VM. 2010 RNA polymerase complexes cooperate to relieve the nucleosomal barrier and evict histones. *Proc. Natl Acad. Sci. USA* **107**, 11 325–11 330. (doi:10.1073/pnas.1001148107)
 38. Weber CM, Ramachandran S, Henikoff S. 2014 Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol. Cell* **53**, 819–830. (doi:10.1016/j.molcel.2014.02.014)
 39. Churchman LS, Weissman JS. 2011 Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**, 368–373. (doi:10.1038/nature09652)
 40. Mayer A, di Iulio J, Maleri S, Eser U, Vierstra J, Reynolds A, Sandstrom R, Stamatoyannopoulos JA, Churchman LS. 2015 Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell* **161**, 541–554. (doi:10.1016/j.cell.2015.03.010)
 41. Kwak H, Fuda NJ, Core LJ, Lis JT. 2013 Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**, 950–953. (doi:10.1126/science.1229386)
 42. Henikoff JG, Belsky JA, Krassovsky K, MacAlpine DM, Henikoff S. 2011 Epigenome characterization at single base-pair resolution. *Proc. Natl Acad. Sci. USA* **108**, 18 318–18 323. (doi:10.1073/pnas.1110731108)
 43. Ramani V, Qiu R, Shendure J. 2019 High sensitivity profiling of chromatin structure by MNase-SSP. *Cell Rep.* **26**, 2465–2476. (doi:10.1016/j.celrep.2019.02.007)
 44. Rhee HS, Bataille AR, Zhang L, Pugh BF. 2014 Subnucleosomal structures and nucleosome asymmetry across a genome. *Cell* **159**, 1377–1388. (doi:10.1016/j.cell.2014.10.054)
 45. Farnung L, Vos SM, Cramer P. 2018 Structure of transcribing RNA polymerase II-nucleosome complex. *Nat. Commun.* **9**, 5432. (doi:10.1038/s41467-018-07870-y)
 46. Kujirai T, Ehara H, Fujino Y, Shirouzu M, Sekine SI, Kurumizaka H. 2018 Structural basis of the nucleosome transition during RNA polymerase II passage. *Science* **362**, 595–598. (doi:10.1126/science.aau9904)
 47. Bilokapic S, Strauss M, Halic M. 2018 Histone octamer rearranges to adapt to DNA unwrapping. *Nat. Struct. Mol. Biol.* **25**, 101–108. (doi:10.1038/s41594-017-0005-5)
 48. Gansauge MT, Meyer M. 2013 Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. *Nat. Protoc.* **8**, 737–748. (doi:10.1038/nprot.2013.038)
 49. Cristiano S *et al.* 2019 Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature* **570**, 385–389. (doi:10.1038/s41586-019-1272-6)
 50. Liu Y *et al.* 2019 Spatial co-fragmentation pattern of cell-free DNA recapitulates *in vivo* chromatin organization and identifies tissues-of-origin. *bioRxiv.* 564773. (doi:10.1101/564773)

51. Markus H *et al.* 2019 Sub-nucleosomal organization in urine cell-free DNA. *bioRxiv*. 696633. (doi:10.1101/696633)
52. Lieberman-Aiden E *et al.* 2009 Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293. (doi:10.1126/science.1181369)
53. Henikoff S, Henikoff JG, Sakai A, Loeb GB, Ahmad K. 2009 Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome Res.* **19**, 460–469. (doi:10.1101/gr.087619.108)
54. Owens N *et al.* 2019 CTCF confers local nucleosome resiliency after DNA replication and during mitosis. *Elife* **8**, e47898. (doi:10.7554/eLife.47898)
55. Carone BR, Hung JH, Hainer SJ, Chou MT, Carone DM, Weng Z, Fazzio TG, Rando OJ. 2014 High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. *Dev. Cell* **30**, 11–22. (doi:10.1016/j.devcel.2014.05.024)
56. Wang J *et al.* 2012 Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res.* **22**, 1798–1812. (doi:10.1101/gr.139105.112)
57. Ulz P *et al.* 2019 Inference of transcription factor binding from cell-free DNA enables tumor subtype prediction and early detection. *Nat. Commun.* **10**, 4666. (doi:10.1038/s41467-019-12714-4)
58. Yevshin I, Sharipov R, Kolmykov S, Kondrakhin Y, Kolpakov F. 2019 GTRD: a database on gene transcription regulation–2019 update. *Nucleic Acids Res.* **47**, D100–D105. (doi:10.1093/nar/gky1128)
59. Chan KC, Zhang J, Hui AB, Wong N, Lau TK, Leung TN, Lo KW, Huang DW, Lo YM. 2004 Size distributions of maternal and fetal DNA in maternal plasma. *Clin. Chem.* **50**, 88–92. (doi:10.1373/clinchem.2003.024893)
60. Sun K *et al.* 2018 Size-tagged preferred ends in maternal plasma DNA shed light on the production mechanism and show utility in noninvasive prenatal testing. *Proc. Natl Acad. Sci. USA* **115**, E5106–E5114. (doi:10.1073/pnas.1804134115)
61. Chabon JJ *et al.* 2020 Integrating genomic features for non-invasive early lung cancer detection. *Nature* **580**, 245–251. (doi:10.1038/s41586-020-2140-0)
62. Mouliere F *et al.* 2018 Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci Transl. Med.* **10**, aat4921. (doi:10.1126/scitranslmed.aat4921)
63. Underhill HR *et al.* 2016 Fragment length of circulating tumor DNA. *PLoS Genet.* **12**, e1006162. (doi:10.1371/journal.pgen.1006162)
64. Gezer U, Ustek D, Yörüker EE, Cakiris A, Abaci N, Leszinski G, Dalay N, Holdenrieder S. 2013 Characterization of H3K9me3- and H4K20me3-associated circulating nucleosomal DNA by high-throughput sequencing in colorectal cancer. *Tumour Biol.* **34**, 329–336. (doi:10.1007/s13277-012-0554-5)
65. Gezer U, Yörüker EE, Keskin M, Kulle CB, Dharuman Y, Holdenrieder S. 2015 Histone methylation marks on circulating nucleosomes as novel blood-based biomarker in colorectal cancer. *Int. J. Mol. Sci.* **16**, 29 654–29 662. (doi:10.3390/ijms161226180)
66. Hoffman MM *et al.* 2013 Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Res.* **41**, 827–841. (doi:10.1093/nar/gks1284)
67. Zentner GE, Henikoff S. 2013 Regulation of nucleosome dynamics by histone modifications. *Nat. Struct. Mol. Biol.* **20**, 259–266. (doi:10.1038/nsmb.2470)
68. Venkatesh S, Workman JL. 2015 Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 178–189. (doi:10.1038/nrm3941)
69. Musselman CA, Lalonde ME, Cote J, Kutateladze TG. 2012 Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.* **19**, 1218–1227. (doi:10.1038/nsmb.2436)
70. 2016 *Chromatin deregulation in cancer*. New York, NY: Cold Spring Harbor Laboratory Press.
71. Morgan MA, Shilatifard A. 2015 Chromatin signatures of cancer. *Genes Dev.* **29**, 238–249. (doi:10.1101/gad.255182.114)
72. Ozdag H *et al.* 2006 Differential expression of selected histone modifier genes in human solid cancers. *BMC Genomics* **7**, 90. (doi:10.1186/1471-2164-7-90)
73. Skulte KA, Phan L, Clark SJ, Taberlay PC. 2014 Chromatin remodeler mutations in human cancers: epigenetic implications. *Epigenomics* **6**, 397–414. (doi:10.2217/epi.14.37)
74. Vaquero A, Scher M, Erdjument-Bromage H, Tempst P, Serrano L, Reinberg D. 2007 SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation. *Nature* **450**, 440–444. (doi:10.1038/nature06268)
75. Fraga MF *et al.* 2005 Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.* **37**, 391–400. (doi:10.1038/ng1531)
76. Hanahan D, Weinberg RA. 2000 The hallmarks of cancer. *Cell* **100**, 57–70. (doi:10.1016/s0092-8674(00)81683-9)
77. Robertson KD. 2005 DNA methylation and human disease. *Nat. Rev. Genet.* **6**, 597–610. (doi:10.1038/nrg1655)
78. Jones PA. 2012 Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492. (doi:10.1038/nrg3230)
79. Suzuki MM, Bird A. 2008 DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* **9**, 465–476. (doi:10.1038/nrg2341)
80. Baron U, Turbachova I, Hellwag A, Eckhardt F, Berlin K, Hoffmuller U, Gardina P, Olek S. 2006 DNA methylation analysis as a tool for cell typing. *Epigenetics* **1**, 55–60. (doi:10.4161/epi.1.1.2643)
81. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. 2012 DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinf.* **13**, 86. (doi:10.1186/1471-2105-13-86)
82. Accomando WP, Wiencke JK, Houseman EA, Nelson HH, Kelsey KT. 2014 Quantitative reconstruction of leukocyte subsets using DNA methylation. *Genome Biol.* **15**, R50. (doi:10.1186/gb-2014-15-3-r50)
83. Esteller M. 2002 CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. *Oncogene* **21**, 5427–5440. (doi:10.1038/sj.onc.1205600)
84. Costello JF *et al.* 2000 Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat. Genet.* **24**, 132–138. (doi:10.1038/72785)
85. McMahon KW, Karunasena E, Ahuja N. 2017 The Roles of DNA methylation in the stages of cancer. *Cancer J.* **23**, 257–261. (doi:10.1097/ppo.0000000000000279)
86. Robertson KD, Jones PA. 2000 DNA methylation: past, present and future directions. *Carcinogenesis* **21**, 461–467. (doi:10.1093/carcin/21.3.461)
87. Usadel H, Brabender J, Danenberg KD, Jeronimo C, Harden S, Engles J, Danenberg PV, Yang S, Sidransky D. 2002 Quantitative adenomatous polyposis coli promoter methylation analysis in tumor tissue, serum, and plasma DNA of patients with lung cancer. *Cancer Res.* **62**, 371–375.
88. Hoon DS, Spugnardi M, Kuo C, Huang SK, Morton DL, Taback B. 2004 Profiling epigenetic inactivation of tumor suppressor genes in tumors and plasma from cutaneous melanoma patients. *Oncogene* **23**, 4014–4022. (doi:10.1038/sj.onc.1207505)
89. Mori T *et al.* 2005 Predictive utility of circulating methylated DNA in serum of melanoma patients receiving biochemotherapy. *J. Clin. Oncol.* **23**, 9351–9358. (doi:10.1200/jco.2005.02.9876)
90. Shen SY *et al.* 2018 Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* **563**, 579–583. (doi:10.1038/s41586-018-0703-0)
91. Liu MC *et al.* Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. *Ann. Oncol.* (doi:10.1016/jannonc.2020.02.011)
92. Hodges C, Kirkland JG, Crabtree GR. 2016 The many roles of BAF (mSWI/SNF) and PBAF complexes in cancer. *Cold Spring Harb. Perspect. Med.* **6**, a026930. (doi:10.1101/cshperspect.a026930)
93. Kim KH, Roberts CW. 2016 Targeting EZH2 in cancer. *Nat. Med.* **22**, 128–134. (doi:10.1038/nm.4036)
94. Sze CC, Shilatifard A. 2016 MLL3/MLL4/COMPASS Family on epigenetic regulation of enhancer function and cancer. *Cold Spring Harb. Perspect. Med.* **6**, a026427. (doi:10.1101/cshperspect.a026427)
95. Wang L, Shilatifard A. 2019 UTX mutations in human cancer. *Cancer Cell* **35**, 168–176. (doi:10.1016/j.ccell.2019.01.001)
96. Zhang W, Xu J. 2017 DNA methyltransferases and their roles in tumorigenesis. *Biomark Res.* **5**, 1. (doi:10.1186/s40364-017-0081-z)
97. Zhao Z, Shilatifard A. 2019 Epigenetic modifications of histones in cancer. *Genome Biol.* **20**, 245. (doi:10.1186/s13059-019-1870-5)