

Review

Next-Generation Sequencing Approaches in Cancer: Where Have They Brought Us and Where Will They Take Us?

Veronique G. LeBlanc ^{1,2} and Marco A. Marra ^{2,3,*}

¹ Genome Science & Technology graduate program, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; E-Mail: vleblanc@bcgsc.ca

² Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, BC V5Z 4S6, Canada

³ Department of Medical Genetics, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

* Author to whom correspondence should be addressed; E-Mail: mmarra@bcgsc.ca; Tel.: +1-604-675-8162; Fax: +1-604-675-8178.

Academic Editors: Camile S. Farah and William Chi-shing Cho

Received: 8 July 2015 / Accepted: 15 September 2015 / Published: 23 September 2015

Abstract: Next-generation sequencing (NGS) technologies and data have revolutionized cancer research and are increasingly being deployed to guide clinicians in treatment decision-making. NGS technologies have allowed us to take an “omics” approach to cancer in order to reveal genomic, transcriptomic, and epigenomic landscapes of individual malignancies. Integrative multi-platform analyses are increasingly used in large-scale projects that aim to fully characterize individual tumours as well as general cancer types and subtypes. In this review, we examine how NGS technologies in particular have contributed to “omics” approaches in cancer research, allowing for large-scale integrative analyses that consider hundreds of tumour samples. These types of studies have provided us with an unprecedented wealth of information, providing the background knowledge needed to make small-scale (including “N of 1”) studies informative and relevant. We also take a look at emerging opportunities provided by NGS and state-of-the-art third-generation sequencing technologies, particularly in the context of translational research. Cancer research and care are currently poised to experience significant progress catalyzed by accessible sequencing technologies that will benefit both clinical- and research-based efforts.

Keywords: cancer; next-generation sequencing; translational research; precision oncology medicine

1. Introduction

Cancer research has witnessed unprecedented advances during the past decade, a good deal of which can be attributed to revolutionary next-generation sequencing (NGS) technologies and the data they provide. Though cancer was first proposed to be a genetic disease early in the 20th century [1], technological limitations restricted progress in identifying the underlying genetic causes of the disease for much of the next century. The past decade has witnessed the advent of NGS and, along with it, a significant increase in our understanding of cancer biology and the cooperative roles of the cancer genome, transcriptome, and epigenome. NGS has made large-scale projects such as The Cancer Genome Atlas (TCGA) [2] and the International Cancer Genome Consortium (ICGC) [3] feasible, providing researchers with multi-platform data for thousands of tumours from a variety of cancer types and subtypes. This has in turn fuelled an expansion of translational research where data obtained from patient samples inform both research and clinical care. Recent case studies demonstrate the utility of whole-genome and -transcriptome sequencing for informing treatment decisions as well as providing valuable research insights. Indeed, these case studies provide excellent examples of the potential of precision (or personalized) oncology medicine, wherein a tumour is characterized at the genomic and/or transcriptomic levels in order to inform treatment. In this review, we will provide an overview of the contributions to cancer research provided by NGS technologies, including rapidly evolving third-generation technologies such as single molecule, real-time (SMRT) sequencing. We will also discuss current translational research efforts and how large-scale, multi-platform integrative analyses have provided us with the background knowledge needed to inform small-scale (including “N of 1”) studies, allowing us to begin taking a precision approach to cancer research and treatment.

2. A Brief History of Sequencing Approaches in Cancer

Throughout the 20th century, groundbreaking and painstaking studies provided insights into the linkages between genes and cancer initiation, development, and aggressiveness. For example, Nowell and Hungerford [4] used cytogenetic methods in 1960 to first describe the Philadelphia chromosome, which underpins chronic myeloid leukemia (CML); however, thirteen years elapsed before the Philadelphia chromosome was identified as the product of a translocation between chromosomes 9 and 22 [5], and the translocation was finally revealed to cause a fusion between the *BCR* and *ABL* genes in 1985 [6]. The fusion protein is a constitutively active tyrosine kinase, and its discovery led to the development of small-molecule tyrosine kinase inhibitors (TKIs) [7]. Though TKI therapy is usually not curative, its introduction revolutionized CML therapy and recent developments continue to improve the standard of care [7].

Technological advances continued to allow for the identification of a wide range of cancer-related genetic anomalies. Cytogenetic techniques, such as fluorescence *in situ* hybridization (FISH) and comparative genome hybridization (CGH), helped identify several additional recurrent chromosomal aberrations and revealed important aspects of cancer biology (e.g., identification of *PIK3CA* as an oncogene in ovarian cancer [8]). In addition, early sequencing technologies developed during the same period by Ray Wu [9–11] and Frederick Sanger [12,13] allowed for the targeted sequencing of specific regions and therefore helped identify recurrent mutations in genes of interest (e.g., hotspot

mutations in the *TP53* gene [14]). In the 1990s, the development of DNA microarrays provided the first opportunity to look at genes on a larger scale. Indeed, oligonucleotide probe arrays (“DNA chips”) provided researchers with a radical new tool capable of providing readouts for many genes in parallel, in contrast to single-gene approaches [15,16]. Chip technology was first applied to human cancer cells in 1996 in an experiment measuring gene expression in the melanoma cell line UACC-903 with and without re-introduction of a normal human chromosome 6, which suppresses the tumorigenic properties of the cell line [17]. The authors were able to show that introduction of the normal chromosome was associated with differential expression of several genes potentially associated with the melanoma phenotype. Soon after, microarrays were used to demonstrate that cancers could be classified based on their gene expression profiles, specifically demonstrating differences between acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL) [18].

In 1990, just five years after Sanger’s chain termination sequencing method was partially automated for the first time [19], the Human Genome Project was launched as a collaboration between genome centres in several countries [20]. A draft sequence was published eleven years later [20], and the project was deemed completed in 2004 [21]. The first phases of the project saw important technological advances, such as the development of improved dye-labeled terminators [22] and thermostable DNA polymerases [23]. Along with the introduction of capillary sequencing [24], these innovations greatly facilitated the completion of the human genome sequence. Advances in automation, such as improved base- and genotype-calling programs [25–27], also helped improve upon the speed of both sequencing and subsequent sequence analysis, setting the stage for further technological advances.

3. Current Sequencing Technologies

3.1. Second-Generation Technologies

The Human Genome Project and its accompanying need for large-scale sequencing approaches and data analysis inspired the creation of NGS methods, which allowed for DNA fragments to be sequenced in a massively parallel fashion. The number of sequence reads obtained using NGS was orders of magnitude higher than that obtained by capillary electrophoresis-based Sanger sequencing; however, this was achieved at the expense of both read length and accuracy. The first commercially available next-generation sequencer provided an ~100-fold increase in throughput relative to contemporary Sanger sequencers; however, read lengths were ~100 base pairs (bp) compared to the ~700 bp read lengths provided by capillary sequencing [28]. Sequencing platforms from different manufacturers were in relatively wide use by 2007 [29], and many improvements in these technologies and the emergence of new ones have subsequently occurred over the past decade. Nevertheless, read lengths provided by second-generation sequencing (SGS, also referred to as short-read sequencing) platforms in use today range from ~100–500 bp [30]. A concise review of these platforms and their sequencing process can be found at [30]. Importantly, NGS technologies also present an exponentially higher computational burden than previous technologies due to their massive increase in throughput [31]. Interpretation of NGS results is particularly complex in the context of cancer research, where a wide range of genomic aberrations can be of interest, and where some alterations may only exist in a subset of tumour cells [32]. Bioinformatics approaches are continuously evolving, and many applications are being developed that help non-bioinformaticians navigate NGS-derived data [33].

3.2. Third-Generation Technologies

3.2.1. SMRT Sequencing

So-called third-generation sequencing (TGS) technologies have recently begun to emerge. This newest generation of sequencing methods is generally composed of technologies that interrogate single DNA molecules instead of clusters of DNA templates, thereby offering several advantages over SGS approaches, such as the elimination of amplification biases [34,35] (Table 1). One of the most widely used TGS technologies is SMRT sequencing, first developed in 2009 [36]. One of the main advantages of SMRT sequencing lies in its ability to produce unusually long read lengths; indeed, average read lengths have been shown to reach 21 kilobases (kb), and continue to improve with the introduction of new reagent kits [37,38]. These read lengths thus allow for the resolution of complex and repetitive genomic regions for a fraction of the time and cost needed for resolution by Sanger sequencing [39]. For instance, high-fidelity long PCR and SMRT sequencing were used to resolve a complex, highly repetitive central exon of the *MUC5AC* gene [40], which has been implicated in colorectal [41] and pancreatic [42] cancer. Advantages provided by SMRT sequencing come at the cost of higher error rates, most often due to insertions and deletions (indels); however, these errors are generally not context-specific, and methods and software that help reduce their impact have been introduced (e.g., the Quiver consensus algorithm [43]). The technology is also biased towards the identification of long fragments, with a recent study showing that novel transcript isoforms less than 300 bp in length identified by short-read sequencing were generally not validated by SMRT sequencing [44]. To address this and other caveats, several groups have shown that a combination of SMRT and short-read sequencing, termed hybrid sequencing, can provide highly accurate sequence results, especially for complex genomic regions [45–47] and transcript isoforms [48–51].

Though SMRT sequencing is often used for the study and assembly of small genomes, such as bacterial genomes [52], the long read lengths are also well suited for sequencing large human cancer-related loci, such as gene fusions products. For instance, a recent study identified *TTYHI-C19MC* fusions driving the expression of the microRNA cluster C19MC in embryonal tumours with multilayered rosettes (ETMRs) [53]. The authors showed that this in turn led to over-expression of an embryonic, brain-specific *DNMT3B* isoform, prompting them to propose a novel model of tumourigenesis for this cancer type. Similarly, two recent studies were able to confirm the presence of kinase domain mutations within *FLT3* genes with activating internal tandem duplications (*FLT3-ITD*) [54,55], which are found in ~20% of AML patients and are associated with poor prognosis. SMRT sequencing has also been used to sequence the entire BCR-ABL1 fusion gene transcript, allowing for the detection of compound mutations and splice isoforms [56]. The authors suggest that such an assay would be beneficial in the clinical setting, where mutations that confer resistance to TKI-based therapy could be readily identified. Though 454 sequencing had previously been used to address this issue, read lengths are not sufficient to cover the entire transcript, introducing the possibility of amplification biases affecting the measurement of mutation frequencies [57,58]. SMRT sequencing has also been applied to the detection of other structural variants, such as deletion and translocation breakpoint determination [59]. Another advantage presented by SMRT sequencing lies in its innate ability to directly detect and differentiate between base

modifications such as 5mC and 5hmC (discussed in more detail below) as a consequence of these base modifications uniquely affecting the output of the method [60].

Table 1. Advantages and limitations of sequencing technologies.

Technology	Advantages	Limitations
Sanger sequencing	<ul style="list-style-type: none"> • Long reads (~700 bp) • High accuracy 	<ul style="list-style-type: none"> • Low throughput
Second-generation sequencing	<ul style="list-style-type: none"> • High throughput 	<ul style="list-style-type: none"> • Short reads (~100–500 bp) • Amplification biases generally occur
Third-generation sequencing	<ul style="list-style-type: none"> • Long reads (average length can reach ~14 kb) • High throughput • No amplification needed • Can detect and differentiate between base modifications • Potential for miniaturization of the technology (nanopore sequencing) 	<ul style="list-style-type: none"> • High error rate • Biased towards long fragments

3.2.2. Nanopore Sequencing

More recently, nanopore-based sequencing technologies have also emerged as promising single-molecule sequencing strategies. Oxford Nanopore, who is leading the development of this technology, released its portable MinION sequencer to a select community of researchers for testing as part of the MinION Access Program (MAP) [61]. The novel sequencing device is USB-powered, allowing for sequencing runs to be performed on a consumer grade computer and thereby greatly increasing both portability and ease of sequencing. In addition, the company promised that the device would cost only \$1000 and provide read lengths orders of magnitude longer than existing technologies. The platform was released in 2014, and preliminary reports suggest that the technology, while promising, requires further improvement [61,62]. Improvements in the chemistry have already led to advances in read quality [62]. The concept of a truly portable high-throughput sequencing platform is attractive in several applications, including fieldwork and point of care diagnostics [61–64], but significant advances are apparently required to fabricate such a device.

4. A Next-Generation “Omics” Approach to Cancer

4.1. The Genome

NGS has provided us with powerful access to the human genome, and has laid the groundwork for the field of cancer genomics. Indeed, given that cancer is fundamentally a genetic disease, NGS was quickly applied to study cancer genomes. In 2008, two groups published full human genome sequences obtained using NGS technologies [65,66]. In the same year, the first cancer genome and the first tumour/normal genome comparison were published in a study that used data from AML tumour cells and matched

normal cells [67]. The study identified ten genes harbouring acquired mutations in the tumour, only two of which had been previously described in AML. The study established that whole-genome sequencing of tumour cells allowed for the identification of cancer-associated mutations, heralding a revolution in cancer research in which similar analyses across an expanding number of tumours and cancer types have been published.

Today, NGS is used in both clinical and research settings. Targeted genetic tests are currently used as diagnostic and prognostic tools in clinical oncology, and more extensive genomic tests seem likely to come into regular use in the near future [68,69]. Targeted cancer panels are advantageous due to their low cost and relatively simple interpretability, and many exist both for specific cancers, such as prostate cancer [70], and for more general application, such as solid tumours [71]. Conversely, whole-genome sequencing (WGS) is more frequently used in the research setting in order to obtain a complete and relatively unbiased view of the genome. Nevertheless, WGS incurs a high sequencing cost and computational burden due to the amount of data produced, and alternate approaches such as whole-exome sequencing (WES) are also used for certain applications [72]. WGS and WES of tumour and matched normal cells allows for the identification of acquired somatic mutations in the tumour, even those that occur at a relatively low frequency if sufficient depth of coverage is achieved [72]. Importantly, WGS requires no *a priori* selection of genes to be profiled, such as the specification of probe sequences often needed for microarray analyses, allowing for the identification of novel alterations. Indeed, WGS provides an unparalleled global view of the genome, allowing for the identification of somatic mutations even in non-coding and unannotated regions of the genome. High-coverage genomic data also allows for the detection of chromosomal rearrangements [73] and for high-resolution identification of copy number variants [74]. The identity and distribution of these and other types of genomic alterations vary widely between cancer types; indeed, patterns of mutations, termed mutational signatures, can be indicative of the underlying mutational processes and are often similar within a cancer type or subtype [75]. Importantly, genetic variants are not generally found uniformly throughout a single tumour; rather, the subclonal distribution of the variants means that they are typically found in 5%–50% of reads obtained by WGS [76]. Though genetic heterogeneity within tumours was recognized before the advent of NGS [77,78], both deep sequencing (*i.e.*, high- or ultra high-coverage) and single-cell sequencing have greatly improved our ability to study this phenomenon [76]. In 2011, Navin *et al.* [79] published the first report of single-cell sequencing in the context of tumour evolution, identifying three distinct subclonal populations within a genetically heterogeneous breast cancer tumour. Analysis of a primary breast tumour and its liver metastasis also suggested that the primary tumour was composed of a single clonal expansion, which also seeded the metastasis. Notably, the sparse coverage obtained by single-cell sequencing only allowed for inferences to be made based on copy number profiles; however, the authors combined data from 100 cells for each tumour, and presented results that were consistent with previous studies suggesting that metastatic cells arise late in tumour development. Though single-cell genomic studies are becoming more frequent, this technology remains to be improved in order to overcome its current limitations, such as whole-genome amplification biases and the prohibitive cost of high sequence coverage [80,81].

The Catalogue of Somatic Mutations in Cancer (COSMIC) is a manually curated database composed of single nucleotide variant (SNV, in both coding and non-coding regions), indel, gene fusion, genome

rearrangement, copy number, and differential expression data from over 12,000 cancer genomes [82]. Approximately half of the COSMIC data is obtained from TCGA and ICGC, and is representative of the power supplied by such large-scale studies to identify recurrent genomic events that participate in cancer initiation and progression. Recent studies using hybrid sequencing have also identified novel structural alterations not previously detected by traditional WGS [51,83]. Further technological advances will undoubtedly continue to reveal additional genomic alterations involved in cancer pathology, and such databases are bound to continue growing.

4.2. The Transcriptome

4.2.1. Gene Expression

In addition to genome characterization, NGS has also been deployed to characterize the cancer transcriptome through RNA sequencing (RNA-seq). This is particularly useful in the context of cancer, a disease characterized by global genomic dysregulation. RNA-seq allows for the quantification of messenger RNAs (mRNAs), providing a measure of gene expression, and can also be used to uncover genome-level alterations as well as novel, disease-associated transcripts and transcript modifications [84]. Indeed, Meyer *et al.* [85] have identified mRNA transcripts containing N^6 -methyladenosine (m^6A), and have shown differences in the levels of this modification in cancer cell lines. Additionally, high RNA-seq coverage allows for the identification of intragenic fusions that may not be identifiable using WGS, including in-frame fusion events that result in the activation of an oncogene [86]. For instance, novel and recurrent kinase gene fusions were recently identified through an analysis of RNA-seq datasets from 12 different cancer types [87]. The study uncovered new fusion partners, such as *TBLIXR1-RET* fusions in thyroid cancer, as well as known fusions in new cancer types, such as *EGFR-SEPT14* fusions in low-grade gliomas, typically found in glioblastoma. The authors suggest that current protocols for targeted genomic profiling of patients would benefit from improvements that allow them to detect such gene fusions, which can in some cases be therapeutically actionable. RNA-seq is also able to detect transcribed somatic mutations [88], as well as isoforms created by alternative splicing events [89,90]. Alternative splicing has been shown to play an important role in tumorigenesis, and can in some cases be a target of therapy (reviewed in [91]). However, short-read sequencing remains limited in its ability to characterize mRNA isoforms, and complex eukaryotic transcriptome analysis will benefit from emerging long-read sequencing technologies that are able to encompass entire transcripts [92]. Indeed, Au *et al.* [50] used hybrid sequencing to characterize human embryonic stem cells (hESCs) and identified thousands of isoforms, including nearly 300 RNAs from novel gene loci.

Nevertheless, gene expression profiling generally remains the most common application of RNA-seq. NGS allows for the quantification of thousands of gene transcripts, and has led to improvements in cancer classification systems, providing enhanced diagnostic, prognostic, and therapeutic criteria. For instance, the IntClust classification system, based on the expression of driver genes, groups breast cancers into clinically and biologically valid subtypes and is better able to explain expression patterns reported by TCGA than the traditional PAM50 classifier [93,94]. Gene expression, either alone or in combination with mutational data, can also be used to investigate spatial and temporal tumour heterogeneity. For instance, several groups have recently analyzed both inter- and intra-tumour heterogeneity in

glioblastoma multiforme (GBM), both of which are highly prevalent and pose serious challenges to the diagnosis and treatment of the disease (for a review, see [95]). Patel *et al.* [96] performed single-cell RNA-seq on 430 cells from five primary GBM tumours and found that each tumour contained a mixture of cells whose profiles correspond to established GBM subtypes. Cells representative of the proneural subtype, for example, were found in all five tumours, regardless of the tumour's dominant subtype. Importantly, the authors suggest that, though established population-level GBM taxonomy has prognostic significance, it does not reflect the transcriptional diversity found within the tumour. Given that this diversity has the potential to confound therapeutic strategies, bulk tumour analyses of highly heterogeneous tumours may not prove sufficient to inform targeted treatment decisions.

In 2014, Tilgner *et al.* [48] published the first personal, allele-specific transcriptome using hybrid sequencing. With this technology, the authors were able to identify novel splicing isoforms, as well as to assess differential allelic expression and isoforms by linking RNA molecules to personal variants on a one-by-one basis. Such in-depth analyses can evidently provide valuable insights into the biology of individual tumours, and may one day be used to guide individualized treatment. However, single-cell genomics and transcriptomics present their own computational challenges, and further advances will be needed before these technologies reach their full potential [80,81]. Given that genomic mutations alone cannot fully explain tumour biology, both individual and large-scale transcriptomic analyses can help improve our understanding of the origin, pathogenesis, and aggressiveness of individual cancers, as well as characterize similarities and differences both within and between cancer types.

4.2.2. Non-Coding RNAs

Though only ~2% of the genome encodes for genes, the Encyclopedia of DNA Elements (ENCODE) has found that ~75% of the genome is transcribed into primary transcripts [97]. Non-coding RNA species such as microRNAs (miRNAs) and long non-coding RNAs (lncRNAs) play important roles in a variety of cellular processes, and have been shown to be widely dysregulated in cancer [98]. MiRNAs are short single-stranded RNA molecules ~22 bases in length that can regulate the expression of target genes through translational repression or mRNA degradation [99]. A single miRNA can regulate the expression of several genes [100], and estimates suggest that up to ~60% of protein-coding genes are targets of miRNA regulation [101]. Unsurprisingly, a large number of miRNAs can thus act as tumour suppressors or oncogenes and exert widespread gene- and pathway-level effects, and their dysregulation is a central event in many cancer types [102]. MicroRNA-31 (miR-31), for instance, has been implicated as a tumour suppressor and an oncogene in different cancer types, and has recently been shown to play a role in cell cycle and epithelial to mesenchymal transition (EMT) regulation in hepatocellular carcinoma [103]. A recent study also showed that miRNAs with abundant expression in diffuse large B-cell lymphoma (DLBCL) regulate the expression of genes involved in metabolism, cell cycle, and protein and chromatin modification [104]. The authors also identified both known and candidate novel miRNAs whose expression is significantly associated with survival. By virtue of their size, miRNAs can be easily isolated and converted to small RNA cDNA libraries, which can then be subjected to massively parallel sequencing (miRNA-seq) [99]. Though microarrays and reverse transcriptase quantitative PCR (RT-qPCR) can be used to quantify miRNA expression, NGS technologies allow for large-scale characterization of both known and novel miRNA species [99]. Indeed, miRNA-seq has been applied

to all TCGA datasets except for GBM [105], demonstrating both the power of the technology and the relevance of miRNA profiling in cancer.

Long non-coding RNAs (lncRNAs), though more poorly characterized than miRNAs, are also increasingly recognized as playing an important role in oncogenesis and cancer pathology. Different lncRNAs can serve diverse functions, such as guiding the site specificity of chromatin-modifying complexes (e.g., XIST and HOTAIR, discussed in more detail below) or acting as regulators of protein signalling pathways involved in carcinogenesis (e.g., lincRNA-21) [106]. The latter instance is particularly interesting in the context of p53, whose transcriptional network includes at least 18 lncRNAs that act in a positive regulatory feedback loop to enhance p53's tumour suppressor activity [107,108]. LincRNA-21, for instance, is activated by p53 and also serves as an effector of p53 signalling through repression of target gene expression, thereby participating in p53-mediated apoptosis [109]. NGS, especially compared to earlier techniques, provides a relatively comprehensive and balanced view of a cell's transcriptome, including previously unannotated species. Traditional RNA-seq includes a library preparation step that selects for polyadenylated transcripts, thereby allowing for the characterization of almost all mRNAs and the majority (~75%) of lncRNAs [110]. Though non-polyadenylated transcript libraries can be constructed [110,111], a recent study has shown that both RNA extraction and library selection methods affect the amount of intronic transcripts included in the library [112]. The authors propose a double selection method that can effectively be used to characterize both polyadenylated and non-polyadenylated transcript fractions. Overall, non-coding RNA studies have greatly benefited from NGS technologies, which allow for the quantification and characterization of both known and novel RNA transcripts. However, ncRNAs are generally expressed to a lower level than mRNAs, and thus require higher coverage to be reliably detected [92]; therefore, ncRNA studies will also benefit from further technological advances that will allow for higher depth of coverage at a lower cost.

4.3. The Epigenome

4.3.1. DNA Modifications

Though cancer is fundamentally a genetic disease, it is also characterized by widespread epigenetic changes, of which aberrant DNA methylation was the first to be recognized and remains one of the most intensively studied. In humans, DNA methylation almost exclusively occurs at the 5' position of a cytosine ring within a CpG dinucleotide [113]. These dinucleotides are found at a lower frequency than would be expected by chance throughout the genome, and typically cluster together in regions termed CpG islands (CGIs) [113]. Feinberg and Vogelstein [114] first reported in 1983 that gene loci could be hypomethylated in tumour cells compared to normal cells. The CpG island methylator phenotype (CIMP), which consists of genome-wide CpG island hypermethylation, was first described more than a decade later [115]. Though originally identified in colorectal cancer, CIMP-positive tumours occur in several cancer types and often constitute a unique subtype (for a review, see [116]). In general, a high proportion of tumours exhibit widespread changes in DNA methylation patterns, most often including global hypomethylation and regional hypermethylation [117]. For instance, the expression of caveolin-1 (*CAVI*), a potential oncogene involved in breast cancer pathogenesis, has been shown to be regulated by differential methylation of CGI shores, which are regions of lower CpG density that

flank CGIs [118]. The authors highlight a negative correlation between the survival rate of estrogen receptor α (ER α)-positive breast cancer patients and *CAVI* expression, suggesting that *CAVI* CGI shore methylation could be used as a prognostic marker for this type of cancer. Studies such as this one reaffirm the usefulness of interrogating the whole genome (or a relatively loosely targeted region of the genome), as relevant changes can occur in regions that were not previously of interest.

Though many methods exist to measure 5-methylcytosine (5mC), the most common involve exposure of DNA to sodium bisulphite, which allows for unmodified cytosine residues to be converted to uracil while 5mC remains stable [119]. Following subsequent PCR amplification, originally unmodified cytosine residues are read as thymine residues, whereas original 5mC residues are read as cytosine residues [119]. Whole-genome bisulphite sequencing (WGBS) is typically considered the gold standard for methylation status interrogation on a genome-wide scale. As the name suggests, WGBS involves sodium bisulphite treatment of DNA followed by NGS, allowing for single-nucleotide resolution of 5mC patterns [120]. Targeted approaches have also been developed to reduce the cost and resources associated with WGBS by isolating a relevant portion of the genome for subsequent sequencing (Table 2). Briefly, targeted approaches include methyl CpG binding domain (MBD)-isolated genome sequencing (MiGS), where MBD precipitation of genomic DNA isolates regions of high methylated CpG density [121]. Reduced representation bisulphite sequencing (RRBS), which makes use of methylation-sensitive restriction enzymes (MSREs) to recover genome-wide CGIs [122], is also frequently used. For a comprehensive review of targeted approaches, see [123]. WGBS has also been successfully applied to single cells, detecting approximately 48% of CpG sites in the human genome [124]. Though the coverage of this and other single-cell WGBS techniques [125] remains to be improved, they provide exciting opportunities for further characterization of epigenetic heterogeneity in cancer cells. WGBS has also recently been applied to archival bone marrow smears from patients with ALL, opening the door to longitudinal epigenetic studies and investigation of temporal tumour heterogeneity [126].

Though 5mC is the most well characterized DNA modification, oxidized variants of 5mC generated by ten eleven translocation (TET) enzymes, namely 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC), can also be found at low frequency throughout the genome [120,127]. Improvements in methods used to quantify and characterize these variants have led to a better understanding and appreciation of their role in cell regulation [120,127]. Conventional bisulphite sequencing cannot distinguish between 5mC and 5hmC, as both variants are resistant to deamination; conversely, 5fC and 5caC are not resistant to bisulphite treatment, but can be converted to resistant residues either by chemical modification or through reduction of 5fC [120]. Methods that allow for relatively quick and inexpensive genome-wide analyses of these variants, such as DNA immunoprecipitation followed by NGS (DIP-seq), chemical labeling, and PCR-based methods are currently being developed and improved (for a review of these methods, see [120]). Single-base pair detection of these variants, however, requires additional resource-intensive methods based on bisulphite sequencing. Oxidative bisulphite sequencing (oxBS-seq) [128] or TET-assisted bisulphite sequencing (TAB-seq) [129] can be used for 5hmC, while chemical modification-assisted bisulphite sequencing (CAB-seq) [130] can be used for 5caC, and fCAB-seq [131] and reduced bisulphite sequencing (redBS-seq) [132] can be used for 5fC (Table 2). For all of these methods, a comparison to WGBS results is needed to differentiate between 5mC and the variant of interest, thereby doubling

the amount of sequencing needed (Plongthongkum *et al.* [120] provide an excellent review of these approaches). Another major caveat of these approaches lies in the read depth needed to produce quantitative results, which is extremely high (usually >1000×) due to the variants' low abundance throughout the genome [133]. Until such drawbacks are addressed, genome-wide, single-nucleotide resolution maps of these variants will continue to be too cost- and resource-intensive to be obtained on a regular basis through conventional sequencing methods. As mentioned above, third-generation sequencing methods are generally able to distinguish between unmethylated cytosine and one or more variants, and hold great promise for the future of genome-wide, single-nucleotide resolution sequencing of DNA modifications [60].

Table 2. NGS-based assays used to study the epigenome.

Acronym	Full Name	Target	Reference(s)
WGBS	Whole-genome bisulphite sequencing	5mC	[134]
MiGS	MBD-isolated genome sequencing	5mC (regions of high methylated CpG density)	[121]
RRBS	Reduced representation bisulphite sequencing	5mC (CpG-rich genomic fragments)	[122]
DIP-seq	DNA immunoprecipitation sequencing	DNA modification of interest	[135–137]
oxBS-seq	Oxidative bisulphite sequencing	5hmC	[128]
TAB-seq	TET-assisted bisulphite sequencing	5hmC	[129]
CAB-seq	Chemical modification-assisted bisulphite sequencing	5caC, 5fC (fCAB-seq)	[130,131]
redBS-seq	Reduced bisulphite sequencing	5fC	[132]
ChIP-seq	Chromatin immunoprecipitation sequencing	Histone modification/histone variant/ DNA-binding protein of interest	[138]
MNase-seq	Micrococcal nuclease sequencing	Nucleosome-associated DNA	[139]
DNase-seq	Deoxyribonuclease sequencing	DNase I hypersensitive sites	[140]
5C	Chromosome conformation capture carbon copy	Spatial distances between genomic regions of interest	[141]

4.3.2. Histone Modifications

In addition to DNA methylation, histone modifications patterns are also often deregulated in cancer, leading to aberrant gene expression and various cancer-related phenotypes. Histone modifications consist of covalent and reversible modifications of core histone proteins, most often by acetylation, phosphorylation, methylation, and/or ubiquitination [142]. The location and composition of the modifications help regulate chromatin accessibility, and therefore influence chromatin-based processes such as DNA transcription and repair [142]. More than a decade ago, Nguyen *et al.* [143] first demonstrated that aberrant histone modifications were likely to play a role in oncogenesis, showing an association between H3K9 hypermethylation and H3K4 hypomethylation and reduced expression of tumour suppressor genes in four cancer cell lines. Indeed, histone modifications are often associated with DNA regulatory elements, and their aberrant localization and abundance can lead to oncogenic

deregulation of gene expression [144]. For instance, ~60% of high-grade paediatric gliomas harbour a recurrent K27M mutation in *H3F3A*, one of two genes that encode the histone H3 variant H3.3 [145]. H3K27 is conserved among all H3 proteins, and can be methylated or acetylated in different contexts: in humans, H3K27ac is found at gene enhancers, while H3K27me3 is enriched at silent gene promoters. The *H3F3A* K27M mutation has been shown to lead to a global reduction in H3K27me2 and H3K27me3 levels and to local increases in H3K27me3 and EZH2 (the catalytic subunit of H3K27 methyltransferase), the sum of which results in increased expression of cancer-associated genes [145]. Histone modifications also function in cooperation with DNA methylation, and each can influence the location and composition of the other. For instance, a recent report used data from the Roadmap Epigenomics Consortium project [146] and ENCODE [147] to identify recurrent gene promoter hypomethylation events in GBM that co-occur with H3K4me3 events [148]. They showed that the transcription of affected genes was altered and that, in at least one case, the modifications led to the expression of an oncogenic protein.

The introduction of the chromatin immunoprecipitation (ChIP) method in 1999 [149] represented a turning point for histone modification studies by allowing for the isolation of specific histone modifications along with the fragment of DNA to which the histone is bound. Today, ChIP is combined with NGS (ChIP-seq) in order to sequence these DNA fragments and map them back to a reference genome, providing a genome-wide view of sequences bound by the histone modification of interest [138]. Though ChIP-seq reduces biases associated with earlier array- or PCR-based techniques, its resolution is relatively low due to the large DNA fragments produced by random shearing of the DNA crosslinked with the immunoprecipitated protein of interest [138]. The method is also sensitive to several factors such as antibody specificity and DNA shearing, the latter of which is not generally highly reproducible. ChIP-exo is a modification of the method that uses an exonuclease enzyme to digest the DNA fragments, thereby improving the resolution with which binding sites can be identified [150,151]. An updated ChIP-exo protocol, named ChIP-nexus (chromatin immunoprecipitation experiments with nucleotide resolution through exonuclease, unique barcode, and single ligation), was also recently published that reduces the amount of starting material needed compared to ChIP-exo [152]. Importantly, ChIP-seq can also be used to map the binding sites of DNA-binding proteins (e.g., transcription factors), as well as the localization of histone variants (e.g., H2A.Z) in addition to histone modifications, all of which participate in the regulation of gene expression [153].

4.3.3. Chromatin Organization

Additional aspects of chromatin organization can also have significant effects on the regulation of gene expression. For instance, euchromatin comprises regions of the genome that are actively transcribed and lightly stained by cytogenetic techniques, while heterochromatin is composed of highly condensed and stainable material associated with silent genes [154]. In 2011, two groups reported that long-range DNA interactions and higher-order chromatin architecture predict the landscape of somatic copy-number alterations (SCNAs) in cancer genomes [155,156]. Indeed, a recent study recently found that chromatin features from the cell of origin are better predictors of a tumour's mutation profile than chromatin features from matched cancer cell lines, indicating that chromatin organization in the cell of origin plays an important role in shaping the future mutational profile of tumour cells [157]. An additional report also showed that, in three colorectal adenocarcinomas, approximately one third of the genome

exhibited substantial DNA demethylation compared to normal mucosal samples, and that these regions corresponded to higher-order domains that are indicative of large-scale chromatin decondensation [158]. Importantly, some genes within these regions showed hypervariability of expression within the tumours, which the authors suggested could contribute to tumour cell heterogeneity and associated resistance mechanisms. Though higher-order chromatin organization has been studied less extensively than genomic, transcriptomic, and even other epigenetic aberrations in the context of cancer, it is becoming evident that the three-dimensional structure of chromatin can play an important role in cancer initiation and progression. For a review of higher-order chromatin organization studies, see [159].

DNA digestion with micrococcal nuclease (MNase) or deoxyribonuclease (DNase) followed by sequencing of the resulting fragments can be used as a measure of nucleosome occupancy and chromatin accessibility, respectively [160]. Nucleosome-associated DNA is protected from MNase digestion, which preferentially cleaves “linker” DNA positioned between nucleosomes; therefore, sequencing of undigested DNA allows for the construction of a genome-wide nucleosome map [139,161]. Conversely, DNase I hypersensitive sites are nucleosome-depleted regions of DNA that are more susceptible to digestion by DNase I. Attachment of biotinylated linkers to DNase I-digested fragments thus enables their isolation and subsequent sequencing, thereby allowing for mapping of DNase I hypersensitive sites (*i.e.*, regions of high chromatin accessibility) [140,162]. The three dimensional structure of chromatin within the nucleus has also been shown to influence gene expression and can be probed using a variety of chromosome conformation capture (3C)-based assays. These assays make use of formaldehyde fixation to cross-link sections of the genome that are in physical contact through DNA-bound proteins [163]. The frequency and location of these interactions can then be assessed to provide information about general nuclear organization and chromosome conformation. Chromosome conformation capture carbon copy (5C), is the first 3C-based assay to make use of NGS, and allows for high-throughput studies that measure spatial distances between many regions of the genome [141].

4.3.4. Non-Coding RNAs

Non-coding RNAs, as mentioned above, are increasingly understood to play an important role in epigenetic regulation. Several lncRNA species in particular have been shown to guide epigenetic complexes to their genomic targets, and the deregulation of one or more lncRNA species is involved in many types of cancer. Elevated HOTAIR levels in particular have been reported in several cancers, and have been linked to metastasis and poor prognosis due to its regulatory role in invasive and proliferative phenotypes [164,165]. HOTAIR is involved in the regulation and targeting of Polycomb repressive complex 2 (PRC2), which is a histone methyltransferase that targets H3K27 [166]. HOTAIR is also involved in the regulation of other signalling pathways, such as the Wnt/ β -catenin pathway [167], and its aberrantly high expression has been shown to trigger EMT and to help cancer cells acquire a stemness phenotype [168]. Similarly, some miRNAs target epigenetic regulators, their aberrant expression thus leading to further epigenetic dysregulation. For instance, a recent report shows that a novel miRNA over-expressed in DLBCL (miR-10393-3p) is significantly anti-correlated with the expression of *KMT2D* and *EP300*, two genes that are involved in chromatin regulation and that are recurrently mutated in non-Hodgkin lymphoma (NHL) [104]. Several reviews further cover the role

of lncRNAs, miRNAs, and other ncRNAs in cancer, many of which function through regulation of epigenetic mechanisms [102,106,169].

4.4. Integrative Analyses

Though individual “omics” studies are informative on their own, integrative analyses that incorporate genomic, transcriptomic, and/or epigenetic information can be extremely powerful. For instance, integration of DNA and RNA sequencing data has recently been shown to help increase the sensitivity of variant detection, especially in low purity tumours [170]. Indeed, the software developed by the authors was able to detect up to >100% increases in mutation rates for certain genes in breast and lung cancers compared to studies published by TCGA. Within the cohorts used for the study, the new method identified a higher number of patients with mutations in genes that could theoretically be therapeutically targeted (e.g., *PIK3CA* and *ERBB2*), suggesting that integration of RNA sequencing information and the resulting increased mutation detection sensitivity may be relevant in a clinical setting when determining optimal treatment options. Classification efforts have also greatly benefitted from genomic and transcriptomic analyses and from studies that take multi-platform data into account. TCGA’s comprehensive characterization of gastric adenocarcinoma, for example, revealed four subtypes based on mutation, copy number, gene expression, and methylation data [171]. Though the subtypes were not found to affect survival, they are characterized by genomic features that may help guide the development of targeted therapies [171,172]. Indeed, molecular classification efforts provide valuable diagnostic, prognostic, and therapeutic insights that can be better used to categorize patients and provide them with optimal treatment options. Medulloblastomas, for instance, were initially described as small round blue cell tumours of the cerebellum; however, the discovery of *hSNF5/INI1* mutations in atypical teratoid/rhabdoid tumours (ATRTs) and C19MC amplifications in ETMRs have led these tumour types to be recognized as different diseases [173]. In addition, true medulloblastoma tumours are now classified into four distinct subtypes that differ in their genomic and transcriptional profiles, demographics, and outcomes, and which can be targeted using different therapies [173].

Tumour heterogeneity studies have also greatly benefitted from integrative NGS analyses. Shah *et al.* [174] were the first to use NGS to interrogate clonal diversity and evolution, and did so at the genomic and transcriptional levels in the context of breast cancer progression. They found that primary tumour samples were more likely than metastatic tumours to reveal true tumour-initiating mutations, given that significant evolution occurred with disease progression and the emergence of metastatic tumours, leading to a higher number of mutations in the latter. They also identified six mutations with intermediate frequencies of 1%–13%, indicating that their presence was restricted to minor subclones and confirming that the tumours were composed of genetically distinct clonal populations. Though it has been established that intra- and inter-tumour heterogeneity have critical implications in diagnosis, prognosis, and treatment of various cancer types, progress remains to be made on how best to address the various problems posed by these realities, such as treatment resistance. NGS of single cells and bulk tumour samples are routinely used to study and characterize tumour heterogeneity [79,96,175–178], and will likely continue to provide valuable insights into evolutionary and resistance mechanisms of specific tumours. Recently, Macaulay *et al.* [179] introduced G&T-seq, a method wherein the genome and the transcriptome of single cells are sequenced in parallel, allowing for inferences to be made regarding the

consequence of genetic abnormalities. Such integrated “omics” analyses of single cells are becoming more accessible, and their implementation will help further elucidate cellular heterogeneity mechanisms in cancer.

Large-scale projects such as TCGA and the ICGC have greatly facilitated integrative analyses by providing multi-platform data for hundreds of tumours across major cancer types. Indeed, sample collection for TCGA ended in 2013, with samples from 11,000 patients across 33 tumour types [180]. The success of the pilot phase, which aimed to characterize the genomic and molecular features of GBM and ovarian carcinoma [181,182], led to an expansion of the project and an increase in the use of NGS technologies. Whole-exome sequences have now been obtained for each participant, and whole-genome sequences are available for a subset of participants [105]. RNA-seq data is also available for each of the 33 tumour types, and miRNA-seq data for all tumour types except GBM. Importantly, most of these data sets are made publicly available. The majority of tumour types (25 of 33) have data from >100 cases, with up to 1098 samples for breast cancer. Integrative, comprehensive analyses have been published by TCGA for fifteen individual cancer type studies to date [171,181–194].

Pan-Cancer Studies

Both TCGA and other groups have also conducted studies that focus on more than one cancer type, such as TCGA’s pan-cancer analysis project that aims to integrate genomic, transcriptomic, and epigenomic data from several cancer types [195]. In 2014, Hoadley *et al.* [196] published an integrative analysis of multi-platform data from the first 12 cancer types studied by TCGA. Their analysis led to a classification of 11 integrated subtypes, some of which present a near perfect relationship with the tissue of origin and some of which are comprised of different cancer types (e.g., the squamous-like subtype is composed of most head and neck squamous cell carcinomas (HNSCCs) and lung squamous cell carcinomas (LUSCs), some bladder adenocarcinomas, and a few lung adenocarcinomas). This reclassification was significantly better associated with disease outcomes than the classic tissue-of-origin taxonomy, and suggested that a substantial amount of patients would benefit from non-standard treatment regimens. Interestingly, bladder cancers were separated into three subtypes in the new taxonomy system, and survival was correlated with subtype membership. These subgroups were subsequently shown to differ in copy number alterations, protein expression (e.g., HER2, Rab25, and markers of EMT), and immune cell signatures. These and other results reported in the study pose important questions as to how cancer as a whole converges onto specific oncogenic mechanisms, and whether a new cancer-wide classification system may provide more relevant diagnostic and prognostic criteria for clinical trial patient selection and overall treatment options.

In addition to improvements in classification, TCGA and similar multi-platform analyses have also helped reveal new aspects of tumour biology. For instance, a recent study reported a novel tumour suppressor role for *ZBTB7A*, whose transcripts are frequently less abundant in solid tumours, and whose protein product acts as a transcriptional repressor of the critical glycolytic genes *GLUT3*, *PFKP*, and *PKM* [197]. Importantly, the authors found that reduced *ZBTB7A* expression was associated with later stages of cancer and poor patient survival across cancer types, but also conferred a higher sensitivity to glycolysis inhibition, which could be exploited to develop new treatment options. Given that some events may occur at a relatively low frequency within cancer types, but be recurrent within, for example,

advanced solid tumours, such a study once again demonstrates the utility of looking beyond traditionally defined cancer types.

Though our understanding of cancer mechanisms has increased dramatically in the past few decades, most cancers, and especially metastatic events, remain difficult to treat and present a high mortality rate. As evidenced by the research done to date, small-scale cancer studies simply do not provide enough information to address the disease in a comprehensive fashion. Large-scale, multi-platform studies, however, provide us with an unprecedented wealth of information and the background knowledge needed to make small-scale studies valuable and informative, as discussed in more detail below.

5. Emerging Opportunities in Translational Research and Personalized Approaches

One of the most valuable aspects of NGS is the fact that it can produce an unparalleled amount of data relatively quickly and inexpensively, allowing for comprehensive projects at a scale that would otherwise not be feasible. A wealth of information can be recovered from a relatively small portion of tissue, providing researchers with unprecedented access to the human cancer genome. Importantly, multi-platform studies can also be conducted, allowing us to examine the interplay between DNA mutations, RNA expression, and epigenomic patterns, obtaining a comprehensive overview of cancer cells. Large-scale projects such as TCGA and the ICGC have already made available data from thousands of tumours across major cancer types, and have helped us refine classification systems as well as our general understanding of cancer biology. This has in turn allowed us to begin taking a more personalized approach to cancer research and treatment, as a single tumour can now be studied in the context of the knowledge we have acquired regarding its cancer type and/or subtype, along with general mechanisms of cancer initiation and progression. Indeed, the concept of precision medicine has been gaining traction in the past years, especially in the field of oncology where a tumour genome can be compared to its matched normal genome in order to characterize the cancer and identify targetable modifications. In-depth characterization of individual tumours is therefore of significant potential benefit to patients, whose treatment can be chosen based on molecular alterations found within their tumour(s). Such studies can also reveal critical aspects of tumour biology, contributing to our more general knowledge of tumour mechanisms and behaviour. For instance, the Personal OncoGenomics (POG) initiative at the British Columbia Cancer Agency is a project that aims to gauge whether in-depth genomic data can be successfully used to guide clinical decision-making while also cataloguing genomic and transcriptomic information for hundreds of cancer patients [198–201]. Though this project is, to our knowledge, the only one that addresses precision oncology measures at the whole-genome scale, other groups have also published similar reports. Published case studies from the POG project and from other institutions, summarized below, are enlightening examples of the benefits afforded by precision approaches, illustrating the broad potential of this approach to multiple tumour types.

Jones *et al.* [202] were the first to use genome-wide analysis of a patient's tumour genome and transcriptome to guide oncology treatment decision-making. Analysis of the primary tumour suggested that it was driven by the *RET* oncogene, offering a rationale for treatment with RET inhibitors. Though the treatment was initially successful, new lesions appeared after seven months. Genomic and transcriptomic analyses of a recurring metastasis detected evidence of activation of the MAPK and AKT pathways, suggesting mutational evolution in response to drug selection and a possible treatment

resistance mechanism. A subsequent POG report described how sequencing of a metastasized sphenoid mass originally considered an undifferentiated squamous cell carcinoma allowed for its re-diagnosis as a *SMARCB1*-negative rhabdoid tumour [198]. The new diagnosis prompted a change in therapy, which initially produced positive results. Similarly, Welch *et al.* [203] presented a case where conventional tests could not determine whether the patient had acute promyelocytic leukemia (PML) or AML with cytogenetics indicative of a poor prognosis. The authors sequenced the patient's genome and were able to identify a cytogenetically cryptic event leading to a *PML-RARA* fusion indicative of PML. The patient, who was originally referred to the centre for treatment with an allogeneic stem cell transplant, was instead treated with all-*trans* retinoic acid (ATRA), and remained in first remission at the time of publication 15 months after presentation. Another group used an in-house-developed panel to identify a BRAF V600E mutation in a patient with intrahepatic cholangiocarcinoma (ICC) [204]. A multidisciplinary tumour board identified the patient as an ideal candidate for dual BRAF and MEK inhibition, the first described instance of this treatment in ICC. At the time of publication, 34 weeks following the beginning of therapy, the patient continued to show response and remained nearly asymptomatic. In the most recent POG study, Sheffield *et al.* [199] reported the characterization of clinically and molecularly dissimilar peritoneal mesothelioma tumours from two patients. This study generated the first whole-genome sequencing data for this rare, relatively poorly understood cancer type. The authors found evidence for the involvement of tumour suppressor genes such as *CDKN2A* and *NF2*, and also revealed novel potential prognostic factors. For instance, one of the patients presented with sarcomatoid histology and multifocal disease, leading to a prediction of poor prognosis; however, the patient responded particularly well to platinum-based chemotherapy, and remained disease-free at the time of publication. The tumour in question was characterized as somatically hypermutated, which has previously been associated with platinum sensitivity in patients with high-grade serous ovarian cancer [205]. This suggests that a hypermutated phenotype of peritoneal mesothelioma might exist and be associated with favourable prognosis. Overall, these studies highlight both the feasibility and the value of precision approaches in oncology, demonstrating significant clinical and research benefits.

The number of studies focusing on patients that respond particularly well to treatment, termed exceptional responders, is also mounting. Tumours that respond exceptionally well to a treatment course, it is reasoned, must harbour unique features that underlie this exceptional response. Such insights may be biologically significant, and may also provide important clues as to how treatment can be improved for normal responders. For instance, tumours with activating *mTOR* mutations [206] or inactivating *TSC1* and *NF2* mutations [207] have both been shown to be sensitive to the mTOR inhibitor everolimus. This indicates that patients who harbour such mutations could be strong candidates for everolimus treatment, and that patients without these mutations may further benefit from alternative therapies. Similarly, Al-Ahmadie *et al.* [208] reported on a patient with metastatic small-cell cancer who achieved a complete and durable response to combined checkpoint kinase 1 (CHK1) inhibition and DNA-damaging chemotherapy. Genomic sequencing revealed a clonal hemizygous mutation in *RAD50*, a member of the Mre11 complex involved in double-stranded break repair. Combined *RAD50* hypomorphism and checkpoint inhibition led to attenuation of both the ATM and ATR axes of the DNA damage response (DDR), thereby conferring extreme sensitivity to DNA-damaging agents. The authors point out that ATM and Mre11 complex mutations occur in various cancer types, providing the opportunity to manipulate this

sensitizing interaction in various contexts for clinical benefit. In fact, Weber and Ryan [209] suggested that ATM and ATR can serve as therapeutic targets either as a monotherapy in cancers with DDR defects, or as a combinatorial therapy in other relevant cancers. Outlier responses can thus also reveal synthetic lethal interactions, which are defined as two genes that cause cell death if they are both mutated or therapeutically targeted [210]. Such interactions can be exploited to effectively target and kill cancer cells, as a treatment that targets one gene product will only be effective in cells that also contain a mutation in the second gene (*i.e.*, tumour cells). Large-scale or targeted RNA interference (RNAi) screens can also be used to probe for genes that interact with a gene of interest in a synthetic lethal manner. For instance, a recent study made use of a large-scale RNAi screen to identify the anti-apoptotic *BCL-2* gene as being a synthetic lethal partner of the *IDH1*^{R132H} mutation [211], which is frequently seen in many cancer types, most notably gliomas and AML [212]. Importantly, the authors note that *IDH1/2* mutation status may be an important predictor of *BCL-2* inhibition response, which has previously been shown to vary between patients.

Many exceptional responder studies are “N of 1” studies that focus on a single patient who shows a particularly unusual response to a therapy that otherwise produces poor results. For instance, the study on everolimus sensitivity mentioned above focused on a single patient who achieved a complete and durable response to the drug while enrolled in a phase II clinical trial that ultimately failed to achieve significant progression-free survival [207]. Though “N of 1” studies are often dismissed as anecdotes, they are increasingly recognized as able to provide meaningful clinical and research results, revealing important mechanisms of drug sensitivity and therapeutic resistance [213]. Indeed, the U.S. National Cancer Institute (NCI) recently launched an exceptional responder programme that aims to collect 100 tumour samples in an effort to identify unifying molecular markers [214,215]. This is in line with published POG case studies, which also highlight the relevance of “N of 1” studies in terms of both patient benefit and research-advancing biological inferences. Furthermore, a new journal named *Molecular Case Studies* is being launched in October 2015 that will focus on genomic and molecular analyses for precision medicine, both in cancer and in other disease fields [216]. Deep sequencing is of particular consequence in these studies, as it is able to provide significant data from a single sample and allows for comparisons across an ever-expanding set of sequenced cancers.

The continually decreasing cost of NGS is compatible with the notion that patient genomes, transcriptomes, and/or epigenomes may soon be routinely obtained and become accessible for targeted studies, possibly through additional projects similar to the POG initiative and the exceptional responder programme. Decreasing sequencing costs are of particular consequence in regards to sequencing of DNA modifications such as 5mC, which require a higher sequencing depth than regular genome sequencing. We can therefore anticipate continued growth in NGS-driven multi-platform integrative analyses of cancer genomes, which will contribute to an improved understanding of cancer initiation and progression mechanisms, thereby revealing new therapeutic targets and enhancing treatment options.

6. Conclusions

The major technological advancements that have been achieved over the past decade, and the use of these to address fundamental problems in cancer research, present an encouraging outlook for the continued application of NGS in an expanding number of different clinical and research settings. Our

knowledge of cancer is increasing, leading to options for more effective screening, prevention, and treatment options. NGS technologies have revolutionized the way we study cancer, and have played a significant role in our increased understanding of the disease. Large-scale single- and multi-platform studies in particular have shed light on multiple aspects of cancer biology, from identifying new relevant mutations to characterizing the interplay between the genome, the transcriptome, and the epigenome within the context of different cancer types. This has in turn led to enhanced classification systems and therefore diagnostic and prognostic knowledge, as well as the consideration of targeted therapies that may be deployed in patients with specific molecular alterations. The knowledge generated provides us with the background needed for small-scale studies, including those with an N of 1, to be thoroughly relevant and informative. This is exemplified in published case studies that have provided evidence for the feasibility and relevance of using whole-genome characterization to guide treatment decisions. Importantly, these studies have also shed light on the biology of rare cancer types such as peritoneal mesothelioma, for which genome-wide characterization had not yet been performed. Expectations for the next decade and beyond are high, and it is hoped that the continued evolution of sequencing technologies will help us continue to more successfully manage cancers through more effective screening and monitoring techniques, and to improve therapeutic and curative treatments through targeted treatment.

Acknowledgments

V.G.L. is a student in the Genome Science and Technology program at the University of British Columbia and is supported by a Canadian Institutes of Health Research Master's Award. M.A.M. holds the University of British Columbia Canada Research Chair in Genome Science. The BC Cancer Agency Genome Sciences Centre gratefully acknowledges support from the BC Cancer Foundation, Genome Canada, Genome British Columbia, the Cancer Research Society, the University of British Columbia, and the Leukemia and Lymphoma Society of Canada.

Author Contributions

V.G.L. wrote the initial draft. M.A.M. provided guidance and assisted in writing and editing the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Boveri, T. Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. *J. Cell Sci.* **2008**, *121*, 1–84. [[CrossRef](#)] [[PubMed](#)]
2. The Cancer Genome Atlas. Available online: <http://cancergenome.nih.gov/> (accessed on 9 September 2015).
3. International Cancer Genome Consortium. Available online: <https://icgc.org/> (accessed on 9 September 2015).

4. Nowell, P.C.; Hungerford, D.A. Chromosome studies on normal and leukemic human leukocytes. *J. Natl. Cancer Inst.* **1960**, *25*, 85–109. [[PubMed](#)]
5. Rowley, J.D. A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* **1973**, *243*, 290–293. [[CrossRef](#)] [[PubMed](#)]
6. Heisterkamp, N.; Stam, K.; Groffen, J.; de Klein, A.; Grosveld, G. Structural organization of the bcr gene and its role in the Ph' translocation. *Nature* **1985**, *315*, 758–761. [[CrossRef](#)] [[PubMed](#)]
7. O'Hare, T.; Zabriskie, M.S.; Eiring, A.M.; Deininger, M.W. Pushing the limits of targeted therapy in chronic myeloid leukaemia. *Nat. Rev. Cancer* **2012**, *12*, 513–526. [[CrossRef](#)] [[PubMed](#)]
8. Shayesteh, L.; Lu, Y.; Kuo, W.L.; Baldocchi, R.; Godfrey, T.; Collins, C.; Pinkel, D.; Powell, B.; Mills, G.B.; Gray, J.W. PIK3CA is implicated as an oncogene in ovarian cancer. *Nat. Genet.* **1999**, *21*, 99–102. [[CrossRef](#)] [[PubMed](#)]
9. Wu, R.; Kaiser, A.D. Structure and base sequence in the cohesive ends of bacteriophage lambda DNA. *J. Mol. Biol.* **1968**, *35*, 523–537. [[CrossRef](#)]
10. Wu, R. Nucleotide sequence analysis of DNA. I. Partial sequence of the cohesive ends of bacteriophage lambda and 186 DNA. *J. Mol. Biol.* **1970**, *51*, 501–521. [[CrossRef](#)]
11. Wu, R.; Taylor, E. Nucleotide sequence analysis of DNA. II. Complete nucleotide sequence of the cohesive ends of bacteriophage lambda DNA. *J. Mol. Biol.* **1971**, *57*, 491–511. [[CrossRef](#)]
12. Sanger, F.; Donelson, J.E.; Coulson, A.R.; Kössel, H.; Fischer, D. Use of DNA polymerase I primed by a synthetic oligonucleotide to determine a nucleotide sequence in phage f1 DNA. *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 1209–1213. [[CrossRef](#)] [[PubMed](#)]
13. Sanger, F.; Coulson, A.R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **1975**, *94*, 441–448. [[CrossRef](#)]
14. Nigro, J.M.; Baker, S.J.; Preisinger, A.C.; Jessup, J.M.; Hostetter, R.; Cleary, K.; Bigner, S.H.; Davidson, N.; Baylin, S.; Devilee, P.; *et al.* Mutations in the p53 gene occur in diverse human tumour types. *Nature* **1989**, *342*, 705–708. [[CrossRef](#)] [[PubMed](#)]
15. Pease, A.C.; Solas, D.; Sullivan, E.J.; Cronin, M.T.; Holmes, C.P.; Fodor, S.P. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5022–5026. [[CrossRef](#)] [[PubMed](#)]
16. Schena, M.; Shalon, D.; Davis, R.W.; Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **1995**, *270*, 467–470. [[CrossRef](#)] [[PubMed](#)]
17. DeRisi, J.; Penland, L.; Brown, P.O.; Bittner, M.L.; Meltzer, P.S.; Ray, M.; Chen, Y.; Su, Y.A.; Trent, J.M. Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* **1996**, *14*, 457–460. [[PubMed](#)]
18. Golub, T.R.; Slonim, D.K.; Tamayo, P.; Huard, C.; Gaasenbeek, M.; Mesirov, J.P.; Coller, H.; Loh, M.L.; Downing, J.R.; Caligiuri, M.A.; *et al.* Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science* **1999**, *286*, 531–537. [[CrossRef](#)] [[PubMed](#)]

19. Smith, L.M.; Sanders, J.Z.; Kaiser, R.J.; Hughes, P.; Dodd, C.; Connell, C.R.; Heiner, C.; Kent, S.B.; Hood, L.E. Fluorescence detection in automated DNA sequence analysis. *Nature* **1986**, *321*, 674–679. [[CrossRef](#)] [[PubMed](#)]
20. Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; *et al.* Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*, 860–921. [[CrossRef](#)] [[PubMed](#)]
21. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **2004**, *431*, 931–945. [[CrossRef](#)] [[PubMed](#)]
22. Rosenblum, B.B.; Lee, L.G.; Spurgeon, S.L.; Khan, S.H.; Menchen, S.M.; Heiner, C.R.; Chen, S.M. New dye-labeled terminators for improved DNA sequencing patterns. *Nucleic Acids Res.* **1997**, *25*, 4500–4504. [[CrossRef](#)] [[PubMed](#)]
23. Reeve, M.A.; Fuller, C.W. A novel thermostable polymerase for DNA sequencing. *Nature* **1995**, *376*, 796–797. [[CrossRef](#)] [[PubMed](#)]
24. Marsh, M.; Tu, O.; Dolnik, V.; Roach, D.; Solomon, N.; Bechtol, K.; Smietana, P.; Wang, L.; Li, X.; Cartwright, P.; *et al.* High-throughput DNA sequencing on a capillary array electrophoresis system. *J. Capillary Electrophor.* **1997**, *4*, 83–89. [[PubMed](#)]
25. Ewing, B.; Hillier, L.; Wendl, M.C.; Green, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **1998**, *8*, 175–185. [[CrossRef](#)] [[PubMed](#)]
26. Ewing, B.; Green, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **1998**, *8*, 186–194. [[CrossRef](#)] [[PubMed](#)]
27. Nickerson, D.A.; Tobe, V.O.; Taylor, S.L. PolyPhred: Automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res.* **1997**, *25*, 2745–2751. [[CrossRef](#)] [[PubMed](#)]
28. Margulies, M.; Egholm, M.; Altman, W.E.; Attiya, S.; Bader, J.S.; Bemben, L.A.; Berka, J.; Braverman, M.S.; Chen, Y.-J.; Chen, Z.; *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **2005**, *437*, 376–380. [[CrossRef](#)] [[PubMed](#)]
29. Mardis, E.R. Next-generation DNA sequencing methods. *Annu. Rev. Genomics Hum. Genet.* **2008**, *9*, 387–402. [[CrossRef](#)] [[PubMed](#)]
30. Reuter, J.A.; Spacek, D.V.; Snyder, M.P. High-throughput sequencing technologies. *Mol. Cell* **2015**, *58*, 586–597. [[CrossRef](#)] [[PubMed](#)]
31. Mardis, E.R. A decade’s perspective on DNA sequencing technology. *Nature* **2011**, *470*, 198–203. [[CrossRef](#)] [[PubMed](#)]
32. Ding, L.; Wendl, M.C.; McMichael, J.F.; Raphael, B.J. Expanding the computational toolbox for mining cancer genomes. *Nat. Rev. Genet.* **2014**, *15*, 556–570. [[CrossRef](#)] [[PubMed](#)]
33. Kouskoumvekaki, I.; Shublaq, N.; Brunak, S. Facilitating the use of large-scale biological data and tools in the era of translational bioinformatics. *Brief. Bioinform.* **2014**, *15*, 942–952. [[CrossRef](#)] [[PubMed](#)]
34. Schadt, E.E.; Turner, S.; Kasarskis, A. A window into third-generation sequencing. *Hum. Mol. Genet.* **2010**, *19*, 227–240. [[CrossRef](#)] [[PubMed](#)]
35. Roberts, R.J.; Carneiro, M.O.; Schatz, M.C. The advantages of SMRT sequencing. *Genome Biol.* **2013**, *14*, 401–404. [[CrossRef](#)] [[PubMed](#)]

36. Eid, J.; Fehr, A.; Gray, J.; Luong, K.; Lyle, J.; Otto, G.; Peluso, P.; Rank, D.; Baybayan, P.; Bettman, B.; *et al.* Real-time DNA sequencing from single polymerase molecules. *Science* **2009**, *323*, 133–138. [[CrossRef](#)] [[PubMed](#)]
37. Kim, K.E.; Peluso, P.; Babayan, P.; Yeadon, P.J.; Yu, C.; Fisher, W.W.; Chin, C.S.; Rapicavoli, N.A.; Rank, D.R.; Li, J.; *et al.* Long-read, whole-genome shotgun sequence data for five model organisms. *Sci Data* **2014**, *1*. [[CrossRef](#)] [[PubMed](#)]
38. PacBio Blog. New Chemistry Boosts Average Read Length to 10 kb–15 kb for PacBio RS II. Available online: <http://blog.pacificbiosciences.com/2014/10/new-chemistry-boosts-average-read.html> (accessed on 14 August 2015).
39. Huddleston, J.; Ranade, S.; Malig, M.; Antonacci, F.; Chaisson, M.; Hon, L.; Sudmant, P.H.; Graves, T.A.; Alkan, C.; Dennis, M.Y.; *et al.* Reconstructing complex regions of genomes using long-read sequencing technology. *Genome Res.* **2014**, *24*, 688–696. [[CrossRef](#)] [[PubMed](#)]
40. Guo, X.; Zheng, S.; Dang, H.; Pace, R.G.; Stonebraker, J.R.; Jones, C.D.; Boellmann, F.; Yuan, G.; Haridass, P.; Fedrigo, O.; *et al.* Genome reference and sequence variation in the large repetitive central exon of human MUC5AC. *Am. J. Respir. Cell Mol. Biol.* **2014**, *50*, 223–232. [[PubMed](#)]
41. Walsh, M.D.; Clendenning, M.; Williamson, E.; Pearson, S.A.; Walters, R.J.; Nagler, B.; Packenas, D.; Win, A.K.; Hopper, J.L.; Jenkins, M.A.; *et al.* Expression of MUC2, MUC5AC, MUC5B, and MUC6 mucins in colorectal cancers and their association with the CpG island methylator phenotype. *Mod. Pathol.* **2013**, *26*, 1642–1656. [[CrossRef](#)] [[PubMed](#)]
42. Liu, D.; Chang, C.H.; Gold, D.V.; Goldenberg, D.M. Identification of PAM4 (clivatuzumab)-reactive epitope on MUC5AC: A promising biomarker and therapeutic target for pancreatic cancer. *Oncotarget* **2015**, *6*, 4274–4285. [[CrossRef](#)] [[PubMed](#)]
43. Chin, C.S.; Alexander, D.H.; Marks, P.; Klammer, A.A.; Drake, J.; Heiner, C.; Clum, A.; Copeland, A.; Huddleston, J.; Eichler, E.E.; *et al.* Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* **2013**, *10*, 563–569. [[CrossRef](#)] [[PubMed](#)]
44. Chen, L.; Kostadima, M.; Martens, J.H.; Canu, G.; Garcia, S.P.; Turro, E.; Downes, K.; Macaulay, I.C.; Bielczyk-Maczynska, E.; Coe, S.; *et al.* Transcriptional diversity during lineage commitment of human blood progenitors. *Science* **2014**, *345*. [[CrossRef](#)] [[PubMed](#)]
45. Koren, S.; Schatz, M.C.; Walenz, B.P.; Martin, J.; Howard, J.T.; Ganapathy, G.; Wang, Z.; Rasko, D.A.; McCombie, W.R.; Jarvis, E.D.; *et al.* Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nat. Biotechnol.* **2012**, *30*, 693–700. [[CrossRef](#)] [[PubMed](#)]
46. Ritz, A.; Bashir, A.; Sindi, S.; Hsu, D.; Hajirasouliha, I.; Raphael, B.J. Characterization of structural variants with single molecule and hybrid sequencing approaches. *Bioinformatics* **2014**, *30*, 3458–3466. [[CrossRef](#)] [[PubMed](#)]
47. Doi, K.; Monjo, T.; Hoang, P.H.; Yoshimura, J.; Yurino, H.; Mitsui, J.; Ishiura, H.; Takahashi, Y.; Ichikawa, Y.; Goto, J.; *et al.* Rapid detection of expanded short tandem repeats in personal genomics using hybrid sequencing. *Bioinformatics* **2014**, *30*, 815–822. [[CrossRef](#)] [[PubMed](#)]
48. Tilgner, H.; Grubert, F.; Sharon, D.; Snyder, M.P. Defining a personal, allele-specific, and single-molecule long-read transcriptome. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9869–9874. [[CrossRef](#)] [[PubMed](#)]

49. Weirather, J.L.; Afshar, P.T.; Clark, T.A.; Tseng, E.; Powers, L.S.; Underwood, J.G.; Zabner, J.; Korlach, J.; Wong, W.H.; Au, K.F. Characterization of fusion genes and the significantly expressed fusion isoforms in breast cancer by hybrid sequencing. *Nucleic Acids Res.* **2015**. [[CrossRef](#)] [[PubMed](#)]
50. Au, K.F.; Sebastiano, V.; Afshar, P.T.; Durruthy, J.D.; Lee, L.; Williams, B.A.; van Bakel, H.; Schadt, E.E.; Reijo-Pera, R.A.; Underwood, J.G.; *et al.* Characterization of the human ESC transcriptome by hybrid sequencing. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4821–4830. [[CrossRef](#)] [[PubMed](#)]
51. Pendleton, M.; Sebra, R.; Pang, A.W.; Ummat, A.; Franzen, O.; Rausch, T.; Stutz, A.M.; Stedman, W.; Anantharaman, T.; Hastie, A.; *et al.* Assembly and diploid architecture of an individual human genome via single-molecule technologies. *Nat. Methods* **2015**, *12*, 780–786. [[CrossRef](#)] [[PubMed](#)]
52. Fomenkov, A.; Lunnen, K.D.; Zhu, Z.; Anton, B.P.; Wilson, G.G.; Vincze, T.; Roberts, R.J. Complete genome sequence and methylome analysis of bacillus strain x1. *Genome Announc.* **2015**, *3*, e01593–e015914. [[CrossRef](#)] [[PubMed](#)]
53. Kleinman, C.L.; Gerges, N.; Papillon-Cavanagh, S.; Sin-Chan, P.; Pramatarova, A.; Quang, D.A.; Adoue, V.; Busche, S.; Caron, M.; Djambazian, H.; *et al.* Fusion of TTYH1 with the C19MC microRNA cluster drives expression of a brain-specific DNMT3B isoform in the embryonal brain tumor ETMR. *Nat. Genet.* **2014**, *46*, 39–44. [[CrossRef](#)] [[PubMed](#)]
54. Smith, C.C.; Wang, Q.; Chin, C.S.; Salerno, S.; Damon, L.E.; Levis, M.J.; Perl, A.E.; Travers, K.J.; Wang, S.; Hunt, J.P.; *et al.* Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia. *Nature* **2012**, *485*, 260–263. [[CrossRef](#)] [[PubMed](#)]
55. Smith, C.C.; Zhang, C.; Lin, K.C.; Lasater, E.A.; Zhang, Y.; Massi, E.; Damon, L.E.; Pendleton, M.; Bashir, A.; Sebra, R.; *et al.* Characterizing and Overriding the Structural Mechanism of the Quizartinib-Resistant FLT3 “Gatekeeper” F691L Mutation with PLX3397. *Cancer Discov.* **2015**, *5*, 668–679. [[CrossRef](#)] [[PubMed](#)]
56. Cavelier, L.; Ameer, A.; Haggqvist, S.; Hoijer, I.; Cahill, N.; Olsson-Stromberg, U.; Hermanson, M. Clonal distribution of BCR-ABL1 mutations and splice isoforms by single-molecule long-read RNA sequencing. *BMC Cancer* **2015**, *15*, 41–45. [[CrossRef](#)] [[PubMed](#)]
57. Soverini, S.; De Benedittis, C.; Machova Polakova, K.; Brouckova, A.; Horner, D.; Iacono, M.; Castagnetti, F.; Gugliotta, G.; Palandri, F.; Papayannidis, C.; *et al.* Unraveling the complexity of tyrosine kinase inhibitor-resistant populations by ultra-deep sequencing of the BCR-ABL kinase domain. *Blood* **2013**, *122*, 1634–1648. [[CrossRef](#)] [[PubMed](#)]
58. Kastner, R.; Zopf, A.; Preuner, S.; Proll, J.; Niklas, N.; Foskett, P.; Valent, P.; Lion, T.; Gabriel, C. Rapid identification of compound mutations in patients with Philadelphia-positive leukaemias by long-range next generation sequencing. *Eur. J. Cancer* **2014**, *50*, 793–800. [[CrossRef](#)] [[PubMed](#)]
59. Patel, A.; Schwab, R.; Liu, Y.T.; Bafna, V. Amplification and thrifty single-molecule sequencing of recurrent somatic structural variations. *Genome Res* **2014**, *24*, 318–328. [[CrossRef](#)] [[PubMed](#)]

60. Flusberg, B.A.; Webster, D.R.; Lee, J.H.; Travers, K.J.; Olivares, E.C.; Clark, T.A.; Korlach, J.; Turner, S.W. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat. Methods* **2010**, *7*, 461–465. [[CrossRef](#)] [[PubMed](#)]
61. Mikheyev, A.S.; Tin, M.M. A first look at the Oxford Nanopore MinION sequencer. *Mol. Ecol. Resour.* **2014**, *14*, 1097–1102. [[CrossRef](#)] [[PubMed](#)]
62. Quick, J.; Quinlan, A.R.; Loman, N.J. A reference bacterial genome dataset generated on the MinION portable single-molecule nanopore sequencer. *Gigascience* **2014**, *3*, 22. [[CrossRef](#)] [[PubMed](#)]
63. Kilianski, A.; Haas, J.L.; Corriveau, E.J.; Liem, A.T.; Willis, K.L.; Kadavy, D.R.; Rosenzweig, C.N.; Minot, S.S. Bacterial and viral identification and differentiation by amplicon sequencing on the MinION nanopore sequencer. *Gigascience* **2015**, *4*, 11–18. [[CrossRef](#)] [[PubMed](#)]
64. Rusk, N. MinION takes center stage. *Nat. Methods* **2015**, *12*, 12–13. [[CrossRef](#)] [[PubMed](#)]
65. Wheeler, D.A.; Srinivasan, M.; Egholm, M.; Shen, Y.; Chen, L.; McGuire, A.; He, W.; Chen, Y.-J.; Makhijani, V.; Roth, G.T.; *et al.* The complete genome of an individual by massively parallel DNA sequencing. *Nature* **2008**, *452*, 872–876. [[CrossRef](#)] [[PubMed](#)]
66. Bentley, D.R.; Balasubramanian, S.; Swerdlow, H.P.; Smith, G.P.; Milton, J.; Brown, C.G.; Hall, K.P.; Evers, D.J.; Barnes, C.L.; Bignell, H.R.; *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **2008**, *456*, 53–59. [[CrossRef](#)] [[PubMed](#)]
67. Ley, T.J.; Mardis, E.R.; Ding, L.; Fulton, B.; McLellan, M.D.; Chen, K.; Dooling, D.; Dunford-Shore, B.H.; McGrath, S.; Hickenbotham, M.; *et al.* DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* **2008**, *456*, 66–72. [[CrossRef](#)] [[PubMed](#)]
68. Manolio, T.A.; Chisholm, R.L.; Ozenberger, B.; Roden, D.M.; Williams, M.S.; Wilson, R.; Bick, D.; Bottinger, E.P.; Brilliant, M.H.; Eng, C.; *et al.* Implementing genomic medicine in the clinic: The future is here. *Genet. Med.* **2013**, *15*, 258–267. [[CrossRef](#)] [[PubMed](#)]
69. Johansen Taber, K.A.; Dickinson, B.D.; Wilson, M. The promise and challenges of next-generation genome sequencing for clinical care. *JAMA Intern. Med.* **2014**, *174*, 275–280. [[CrossRef](#)] [[PubMed](#)]
70. Leyten, G.H.; Hessels, D.; Smit, F.; Jannink, S.; de Jong, H.; Melchers, W.J.; Cornel, E.B.; de Reijke, T.M.; Vergunst, H.; Kil, P.; *et al.* Identification of a candidate gene panel for the early diagnosis of prostate cancer. *Clin. Cancer Res.* **2015**. [[CrossRef](#)] [[PubMed](#)]
71. Hovelson, D.H.; McDaniel, A.S.; Cani, A.K.; Johnson, B.; Rhodes, K.; Williams, P.D.; Bandla, S.; Bien, G.; Choppa, P.; Hyland, F.; *et al.* Development and validation of a scalable next-generation sequencing system for assessing relevant somatic variants in solid tumors. *Neoplasia* **2015**, *17*, 385–399. [[CrossRef](#)] [[PubMed](#)]
72. Sims, D.; Sudbery, I.; Illott, N.E.; Heger, A.; Ponting, C.P. Sequencing depth and coverage: Key considerations in genomic analyses. *Nat. Rev. Genet.* **2014**, *15*, 121–132. [[CrossRef](#)] [[PubMed](#)]
73. Campbell, P.J.; Stephens, P.J.; Pleasance, E.D.; O'Meara, S.; Li, H.; Santarius, T.; Stebbings, L.A.; Leroy, C.; Edkins, S.; Hardy, C.; *et al.* Identification of somatically acquired rearrangements in

- cancer using genome-wide massively parallel paired-end sequencing. *Nat. Genet.* **2008**, *40*, 722–729. [[CrossRef](#)] [[PubMed](#)]
74. Chiang, D.Y.; Getz, G.; Jaffe, D.B.; O'Kelly, M.J.; Zhao, X.; Carter, S.L.; Russ, C.; Nusbaum, C.; Meyerson, M.; Lander, E.S. High-resolution mapping of copy-number alterations with massively parallel sequencing. *Nat. Methods* **2009**, *6*, 99–103. [[CrossRef](#)] [[PubMed](#)]
75. Helleday, T.; Eshtad, S.; Nik-Zainal, S. Mechanisms underlying mutational signatures in human cancers. *Nat. Rev. Genet.* **2014**, *15*, 585–598. [[CrossRef](#)] [[PubMed](#)]
76. Greaves, M.; Maley, C.C. Clonal evolution in cancer. *Nature* **2012**, *481*, 306–313. [[CrossRef](#)] [[PubMed](#)]
77. Nowell, P.C. The clonal evolution of tumor cell populations. *Science* **1976**, *194*, 23–28. [[CrossRef](#)] [[PubMed](#)]
78. Fidler, I.J. Tumor heterogeneity and the biology of cancer invasion and metastasis. *Cancer Res.* **1978**, *38*, 2651–2660. [[PubMed](#)]
79. Navin, N.; Kendall, J.; Troge, J.; Andrews, P.; Rodgers, L.; McIndoo, J.; Cook, K.; Stepansky, A.; Levy, D.; Esposito, D.; *et al.* Tumour evolution inferred by single-cell sequencing. *Nature* **2011**, *472*, 90–94. [[CrossRef](#)] [[PubMed](#)]
80. Shapiro, E.; Biezuner, T.; Linnarsson, S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat. Rev. Genet.* **2013**, *14*, 618–630. [[CrossRef](#)] [[PubMed](#)]
81. Stegle, O.; Teichmann, S.A.; Marioni, J.C. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* **2015**, *16*, 133–145. [[CrossRef](#)] [[PubMed](#)]
82. Forbes, S.A.; Beare, D.; Gunasekaran, P.; Leung, K.; Bindal, N.; Boutselakis, H.; Ding, M.; Bamford, S.; Cole, C.; Ward, S.; *et al.* COSMIC: Exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **2015**, *43*, 805–811. [[CrossRef](#)] [[PubMed](#)]
83. Chaisson, M.J.; Huddleston, J.; Dennis, M.Y.; Sudmant, P.H.; Malig, M.; Hormozdiari, F.; Antonacci, F.; Surti, U.; Sandstrom, R.; Boitano, M.; *et al.* Resolving the complexity of the human genome using single-molecule sequencing. *Nature* **2015**, *517*, 608–611. [[CrossRef](#)] [[PubMed](#)]
84. Morozova, O.; Hirst, M.; Marra, M.A. Applications of new sequencing technologies for transcriptome analysis. *Annu. Rev. Genomics Hum. Genet.* **2009**, *10*, 135–151. [[CrossRef](#)] [[PubMed](#)]
85. Meyer, K.D.; Saletore, Y.; Zumbo, P.; Elemento, O.; Mason, C.E.; Jaffrey, S.R. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **2012**, *149*, 1635–1646. [[CrossRef](#)] [[PubMed](#)]
86. Maher, C.A.; Kumar-Sinha, C.; Cao, X.; Kalyana-Sundaram, S.; Han, B.; Jing, X.; Sam, L.; Barrette, T.; Palanisamy, N.; Chinnaiyan, A.M. Transcriptome sequencing to detect gene fusions in cancer. *Nature* **2009**, *458*, 97–101. [[CrossRef](#)] [[PubMed](#)]
87. Stransky, N.; Cerami, E.; Schalm, S.; Kim, J.L.; Lengauer, C. The landscape of kinase fusions in cancer. *Nat. Commun.* **2014**, *5*, 4841–4846. [[CrossRef](#)] [[PubMed](#)]
88. Goya, R.; Sun, M.G.; Morin, R.D.; Leung, G.; Ha, G.; Wiegand, K.C.; Senz, J.; Crisan, A.; Marra, M.A.; Hirst, M.; *et al.* SNVMix: Predicting single nucleotide variants from next-generation sequencing of tumors. *Bioinformatics* **2010**, *26*, 730–736. [[CrossRef](#)] [[PubMed](#)]

89. Morin, R.; Bainbridge, M.; Fejes, A.; Hirst, M.; Krzywinski, M.; Pugh, T.; McDonald, H.; Varhol, R.; Jones, S.; Marra, M. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *BioTechniques* **2008**, *45*, 81–94. [[CrossRef](#)] [[PubMed](#)]
90. Sultan, M.; Schulz, M.H.; Richard, H.; Magen, A.; Klingenhoff, A.; Scherf, M.; Seifert, M.; Borodina, T.; Soldatov, A.; Parkhomchuk, D.; *et al.* A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **2008**, *321*, 956–960. [[CrossRef](#)] [[PubMed](#)]
91. Chen, J.; Weiss, W.A. Alternative splicing in cancer: Implications for biology and therapy. *Oncogene* **2015**, *34*, 1–14. [[CrossRef](#)] [[PubMed](#)]
92. Steijger, T.; Abril, J.F.; Engstrom, P.G.; Kokocinski, F.; Hubbard, T.J.; Guigo, R.; Harrow, J.; Bertone, P.; Consortium, R. Assessment of transcript reconstruction methods for RNA-seq. *Nat. Methods* **2013**, *10*, 1177–1184. [[CrossRef](#)] [[PubMed](#)]
93. Dawson, S.-J.; Rueda, O.M.; Aparicio, S.; Caldas, C. A new genome-driven integrated classification of breast cancer and its implications. *EMBO J.* **2013**, *32*, 617–628. [[CrossRef](#)] [[PubMed](#)]
94. Ali, H.R.; Rueda, O.M.; Chin, S.-F.; Curtis, C.; Dunning, M.J.; Aparicio, S.A.; Caldas, C. Genome-driven integrated classification of breast cancer validated in over 7,500 samples. *Genome Biol.* **2014**, *15*. [[CrossRef](#)] [[PubMed](#)]
95. Parker, N.R.; Khong, P.; Parkinson, J.F.; Howell, V.M.; Wheeler, H.R. Molecular heterogeneity in glioblastoma: Potential clinical implications. *Front. Oncol.* **2015**, *5*, 51–59. [[CrossRef](#)] [[PubMed](#)]
96. Patel, A.P.; Tirosh, I.; Trombetta, J.J.; Shalek, A.K.; Gillespie, S.M.; Wakimoto, H.; Cahill, D.P.; Nahed, B.V.; Curry, W.T.; Martuza, R.L.; *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **2014**, *344*, 1396–1401. [[CrossRef](#)] [[PubMed](#)]
97. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; *et al.* Landscape of transcription in human cells. *Nature* **2012**, *489*, 101–108. [[CrossRef](#)]
98. Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **2011**, *12*, 861–874. [[CrossRef](#)] [[PubMed](#)]
99. Pritchard, C.C.; Cheng, H.H.; Tewari, M. MicroRNA profiling: Approaches and considerations. *Nat. Rev. Genet.* **2012**, *13*, 358–369. [[CrossRef](#)] [[PubMed](#)]
100. Lim, L.P.; Lau, N.C.; Garrett-Engele, P.; Grimson, A.; Schelter, J.M.; Castle, J.; Bartel, D.P.; Linsley, P.S.; Johnson, J.M. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **2005**, *433*, 769–773. [[CrossRef](#)] [[PubMed](#)]
101. Friedman, R.C.; Farh, K.K.; Burge, C.B.; Bartel, D.P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **2009**, *19*, 92–105. [[CrossRef](#)] [[PubMed](#)]
102. Farazi, T.A.; Hoell, J.I.; Morozov, P.; Tuschl, T. MicroRNAs in human cancer. *Adv. Exp. Med. Biol.* **2013**, *774*, 1–20. [[PubMed](#)]
103. Kim, H.S.; Lee, K.S.; Bae, H.J.; Eun, J.W.; Shen, Q.; Park, S.J.; Shin, W.C.; Yang, H.D.; Park, M.; Park, W.S.; *et al.* MicroRNA-31 functions as a tumor suppressor by regulating cell

- cycle and epithelial-mesenchymal transition regulatory proteins in liver cancer. *Oncotarget* **2015**, *6*, 8089–8102. [[CrossRef](#)] [[PubMed](#)]
104. Lim, E.L.; Trinh, D.L.; Scott, D.W.; Chu, A.; Krzywinski, M.; Zhao, Y.; Robertson, A.G.; Mungall, A.J.; Schein, J.; Boyle, M.; *et al.* Comprehensive miRNA sequence analysis reveals survival differences in diffuse large B-cell lymphoma patients. *Genome Biol.* **2015**, *16*, 11–19. [[CrossRef](#)] [[PubMed](#)]
105. The Cancer Genome Atlas Data Portal—Data Levels and Data Types. Available online: <https://tcga-data.nci.nih.gov/tcga/tcgaDataType.jsp> (accessed on 25 June 2015).
106. Cheetham, S.W.; Gruhl, F.; Mattick, J.S.; Dinger, M.E. Long noncoding RNAs and the genetics of cancer. *Br. J. Cancer* **2013**, *108*, 2419–2425. [[CrossRef](#)] [[PubMed](#)]
107. Zhang, A.; Xu, M.; Mo, Y.Y. Role of the lncRNA-p53 regulatory network in cancer. *J. Mol. Cell Biol.* **2014**, *6*, 181–191. [[CrossRef](#)] [[PubMed](#)]
108. Sanchez, Y.; Segura, V.; Marin-Bejar, O.; Athie, A.; Marchese, F.P.; Gonzalez, J.; Bujanda, L.; Guo, S.; Matheu, A.; Huarte, M. Genome-wide analysis of the human p53 transcriptional network unveils a lncRNA tumour suppressor signature. *Nat. Commun.* **2014**, *5*, 5811–5813. [[CrossRef](#)] [[PubMed](#)]
109. Huarte, M.; Guttman, M.; Feldser, D.; Garber, M.; Koziol, M.J.; Kenzelmann-Broz, D.; Khalil, A.M.; Zuk, O.; Amit, I.; Rabani, M.; *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **2010**, *142*, 409–419. [[CrossRef](#)] [[PubMed](#)]
110. Yang, L.; Duff, M.O.; Graveley, B.R.; Carmichael, G.G.; Chen, L.L. Genomewide characterization of non-polyadenylated RNAs. *Genome Biol.* **2011**, *12*, 11–14. [[CrossRef](#)] [[PubMed](#)]
111. Zhu, S.; Zhang, X.O.; Yang, L. Panning for Long Noncoding RNAs. *Biomolecules* **2013**, *3*, 226–241. [[CrossRef](#)]
112. Sultan, M.; Amstislavskiy, V.; Risch, T.; Schuette, M.; Dokel, S.; Ralser, M.; Balzereit, D.; Lehrach, H.; Yaspo, M.L. Influence of RNA extraction methods and library selection schemes on RNA-seq data. *BMC Genomics* **2014**, *15*, 671–675. [[CrossRef](#)] [[PubMed](#)]
113. Bird, A.P. CpG-rich islands and the function of DNA methylation. *Nature* **1986**, *321*, 209–213. [[CrossRef](#)] [[PubMed](#)]
114. Feinberg, A.P.; Vogelstein, B. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **1983**, *301*, 89–92. [[CrossRef](#)] [[PubMed](#)]
115. Toyota, M.; Ahuja, N.; Ohe-Toyota, M.; Herman, J.G.; Baylin, S.B.; Issa, J.P. CpG island methylator phenotype in colorectal cancer. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 8681–8686. [[CrossRef](#)]
116. Suzuki, H.; Yamamoto, E.; Maruyama, R.; Niinuma, T.; Kai, M. Biological significance of the CpG island methylator phenotype. *Biochem. Biophys. Res. Commun.* **2014**, *455*, 35–42. [[CrossRef](#)] [[PubMed](#)]
117. Stirzaker, C.; Taberlay, P.C.; Statham, A.L.; Clark, S.J. Mining cancer methylomes: Prospects and challenges. *Trends Genet.* **2014**, *30*, 75–84. [[CrossRef](#)] [[PubMed](#)]

118. Rao, X.; Evans, J.; Chae, H.; Pilrose, J.; Kim, S.; Yan, P.; Huang, R.L.; Lai, H.C.; Lin, H.; Liu, Y.; *et al.* CpG island shore methylation regulates caveolin-1 expression in breast cancer. *Oncogene* **2013**, *32*, 4519–4528. [[CrossRef](#)] [[PubMed](#)]
119. Frommer, M.; McDonald, L.E.; Millar, D.S.; Collis, C.M.; Watt, F.; Grigg, G.W.; Molloy, P.L.; Paul, C.L. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 1827–1831. [[CrossRef](#)] [[PubMed](#)]
120. Plongthongkum, N.; Diep, D.H.; Zhang, K. Advances in the profiling of DNA modifications: Cytosine methylation and beyond. *Nat. Rev. Genet.* **2014**, *15*, 647–661. [[CrossRef](#)] [[PubMed](#)]
121. Serre, D.; Lee, B.H.; Ting, A.H. MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Res.* **2010**, *38*, 391–399. [[CrossRef](#)] [[PubMed](#)]
122. Meissner, A.; Mikkelsen, T.S.; Gu, H.; Wernig, M.; Hanna, J.; Sivachenko, A.; Zhang, X.; Bernstein, B.E.; Nusbaum, C.; Jaffe, D.B.; *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **2008**, *454*, 766–770. [[CrossRef](#)] [[PubMed](#)]
123. Harrison, A.; Parle-McDermott, A. DNA methylation: A timeline of methods and applications. *Front. Genet.* **2011**, *2*, 71–74. [[CrossRef](#)] [[PubMed](#)]
124. Smallwood, S.A.; Lee, H.J.; Angermueller, C.; Krueger, F.; Saadeh, H.; Peat, J.; Andrews, S.R.; Stegle, O.; Reik, W.; Kelsey, G. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods* **2014**, *11*, 817–820. [[CrossRef](#)] [[PubMed](#)]
125. Guo, H.; Zhu, P.; Wu, X.; Li, X.; Wen, L.; Tang, F. Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res.* **2013**, *23*, 2126–2135. [[CrossRef](#)] [[PubMed](#)]
126. Wong, N.C.; Meredith, G.D.; Marnellos, G.; Dudas, M.; Parkinson-Bates, M.; Halemba, M.S.; Chatterton, Z.; Maksimovic, J.; Ashley, D.M.; Mechinaud, F.; *et al.* Paediatric leukaemia DNA methylation profiling using MBD enrichment and SOLiD sequencing on archival bone marrow smears. *Gigascience* **2015**, *4*, 11–16. [[CrossRef](#)]
127. Ito, S.; Shen, L.; Dai, Q.; Wu, S.C.; Collins, L.B.; Swenberg, J.A.; He, C.; Zhang, Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **2011**, *333*, 1300–1303. [[CrossRef](#)] [[PubMed](#)]
128. Booth, M.J.; Branco, M.R.; Ficz, G.; Oxley, D.; Krueger, F.; Reik, W.; Balasubramanian, S. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* **2012**, *336*, 934–937. [[CrossRef](#)] [[PubMed](#)]
129. Yu, M.; Hon, G.C.; Szulwach, K.E.; Song, C.X.; Zhang, L.; Kim, A.; Li, X.; Dai, Q.; Shen, Y.; Park, B.; *et al.* Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* **2012**, *149*, 1368–1380. [[CrossRef](#)] [[PubMed](#)]
130. Lu, X.; Song, C.X.; Szulwach, K.; Wang, Z.; Weidenbacher, P.; Jin, P.; He, C. Chemical modification-assisted bisulfite sequencing (CAB-Seq) for 5-carboxylcytosine detection in DNA. *J. Am. Chem. Soc.* **2013**, *135*, 9315–9317. [[CrossRef](#)] [[PubMed](#)]

131. Song, C.X.; Szulwach, K.E.; Dai, Q.; Fu, Y.; Mao, S.Q.; Lin, L.; Street, C.; Li, Y.; Poidevin, M.; Wu, H.; *et al.* Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. *Cell* **2013**, *153*, 678–691. [[CrossRef](#)] [[PubMed](#)]
132. Booth, M.J.; Marsico, G.; Bachman, M.; Beraldi, D.; Balasubramanian, S. Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nat. Chem.* **2014**, *6*, 435–440. [[CrossRef](#)] [[PubMed](#)]
133. Nestor, C.E.; Reddington, J.P.; Benson, M.; Meehan, R.R. Investigating 5-hydroxymethylcytosine (5hmC): the state of the art. *Methods Mol. Biol.* **2014**, *1094*, 243–258. [[PubMed](#)]
134. Cokus, S.J.; Feng, S.; Zhang, X.; Chen, Z.; Merriman, B.; Haudenschild, C.D.; Pradhan, S.; Nelson, S.F.; Pellegrini, M.; Jacobsen, S.E. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* **2008**, *452*, 215–219. [[CrossRef](#)] [[PubMed](#)]
135. Down, T.A.; Rakyen, V.K.; Turner, D.J.; Flicek, P.; Li, H.; Kulesha, E.; Graf, S.; Johnson, N.; Herrero, J.; Tomazou, E.M.; *et al.* A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat. Biotechnol.* **2008**, *26*, 779–785. [[CrossRef](#)] [[PubMed](#)]
136. Stroud, H.; Feng, S.; Morey Kinney, S.; Pradhan, S.; Jacobsen, S.E. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol.* **2011**, *12*, R54. [[CrossRef](#)] [[PubMed](#)]
137. Pastor, W.A.; Pape, U.J.; Huang, Y.; Henderson, H.R.; Lister, R.; Ko, M.; McLoughlin, E.M.; Brudno, Y.; Mahapatra, S.; Kapranov, P.; *et al.* Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **2011**, *473*, 394–397. [[CrossRef](#)] [[PubMed](#)]
138. Johnson, D.S.; Mortazavi, A.; Myers, R.M.; Wold, B. Genome-wide mapping of *in vivo* protein-DNA interactions. *Science* **2007**, *316*, 1497–1502. [[CrossRef](#)] [[PubMed](#)]
139. Schones, D.E.; Cui, K.; Cuddapah, S.; Roh, T.Y.; Barski, A.; Wang, Z.; Wei, G.; Zhao, K. Dynamic regulation of nucleosome positioning in the human genome. *Cell* **2008**, *132*, 887–898. [[CrossRef](#)] [[PubMed](#)]
140. Boyle, A.P.; Davis, S.; Shulha, H.P.; Meltzer, P.; Margulies, E.H.; Weng, Z.; Furey, T.S.; Crawford, G.E. High-resolution mapping and characterization of open chromatin across the genome. *Cell* **2008**, *132*, 311–322. [[CrossRef](#)] [[PubMed](#)]
141. Dostie, J.; Richmond, T.A.; Arnaout, R.A.; Selzer, R.R.; Lee, W.L.; Honan, T.A.; Rubio, E.D.; Krumm, A.; Lamb, J.; Nusbaum, C.; *et al.* Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements. *Genome Res.* **2006**, *16*, 1299–1309. [[CrossRef](#)] [[PubMed](#)]
142. Kouzarides, T. Chromatin modifications and their function. *Cell* **2007**, *128*, 693–705. [[CrossRef](#)] [[PubMed](#)]
143. Nguyen, C.T.; Weisenberger, D.J.; Velicescu, M.; Gonzales, F.A.; Lin, J.C.; Liang, G.; Jones, P.A. Histone H3-lysine 9 methylation is associated with aberrant gene silencing in cancer cells and is rapidly reversed by 5-aza-2'-deoxycytidine. *Cancer Res.* **2002**, *62*, 6456–6461. [[PubMed](#)]

144. Choe, M.K.; Hong, C.P.; Park, J.; Seo, S.H.; Roh, T.Y. Functional elements demarcated by histone modifications in breast cancer cells. *Biochem. Biophys. Res. Commun.* **2012**, *418*, 475–482. [[CrossRef](#)] [[PubMed](#)]
145. Chan, K.M.; Fang, D.; Gan, H.; Hashizume, R.; Yu, C.; Schroeder, M.; Gupta, N.; Mueller, S.; James, C.D.; Jenkins, R.; *et al.* The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes Dev.* **2013**, *27*, 985–990. [[CrossRef](#)] [[PubMed](#)]
146. Roadmap Epigenomics Consortium; Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **2015**, *518*, 317–330.
147. Encode Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74.
148. Nagarajan, R.P.; Zhang, B.; Bell, R.J.; Johnson, B.E.; Olshen, A.B.; Sundaram, V.; Li, D.; Graham, A.E.; Diaz, A.; Fouse, S.D.; *et al.* Recurrent epimutations activate gene body promoters in primary glioblastoma. *Genome Res.* **2014**, *24*, 761–774. [[CrossRef](#)] [[PubMed](#)]
149. Kuo, M.H.; Allis, C.D. *In vivo* cross-linking and immunoprecipitation for studying dynamic Protein:DNA associations in a chromatin environment. *Methods* **1999**, *19*, 425–433. [[CrossRef](#)] [[PubMed](#)]
150. Rhee, H.S.; Pugh, B.F. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* **2011**, *147*, 1408–1419. [[CrossRef](#)] [[PubMed](#)]
151. Rhee, H.S.; Pugh, B.F. ChIP-exo method for identifying genomic location of DNA-binding proteins with near-single-nucleotide accuracy. *Curr. Protoc. Mol. Biol.* **2012**. [[CrossRef](#)]
152. He, Q.; Johnston, J.; Zeitlinger, J. ChIP-nexus enables improved detection of *in vivo* transcription factor binding footprints. *Nat. Biotechnol.* **2015**, *33*, 395–401. [[CrossRef](#)] [[PubMed](#)]
153. Mundade, R.; Ozer, H.G.; Wei, H.; Prabhu, L.; Lu, T. Role of ChIP-seq in the discovery of transcription factor binding sites, differential gene regulation mechanism, epigenetic marks and beyond. *Cell Cycle* **2014**, *13*, 2847–2852. [[CrossRef](#)] [[PubMed](#)]
154. Fedorova, E.; Zink, D. Nuclear architecture and gene regulation. *Biochim. Biophys. Acta* **2008**, *1783*, 2174–2184. [[CrossRef](#)] [[PubMed](#)]
155. Fudenberg, G.; Getz, G.; Meyerson, M.; Mirny, L.A. High order chromatin architecture shapes the landscape of chromosomal alterations in cancer. *Nat. Biotechnol.* **2011**, *29*, 1109–1113. [[CrossRef](#)] [[PubMed](#)]
156. De, S.; Michor, F. DNA replication timing and long-range DNA interactions predict mutational landscapes of cancer genomes. *Nat. Biotechnol.* **2011**, *29*, 1103–1108. [[CrossRef](#)]
157. Polak, P.; Karlić, R.; Koren, A.; Thurman, R.; Sandstrom, R.; Lawrence, M.S.; Reynolds, A.; Rynes, E.; Vlahoviček, K.; Stamatoyannopoulos, J.A.; *et al.* Cell-of-origin chromatin organization shapes the mutational landscape of cancer. *Nature* **2015**, *518*, 360–364. [[CrossRef](#)] [[PubMed](#)]
158. Hansen, K.D.; Timp, W.; Bravo, H.C.; Sabunciyan, S.; Langmead, B.; McDonald, O.G.; Wen, B.; Wu, H.; Liu, Y.; Diep, D.; *et al.* Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* **2011**, *43*, 768–775. [[CrossRef](#)] [[PubMed](#)]

159. Reddy, K.L.; Feinberg, A.P. Higher order chromatin organization in cancer. *Semin. Cancer Biol.* **2013**, *23*, 109–115. [[CrossRef](#)] [[PubMed](#)]
160. Bell, O.; Tiwari, V.K.; Thoma, N.H.; Schubeler, D. Determinants and dynamics of genome accessibility. *Nat. Rev. Genet.* **2011**, *12*, 554–564. [[CrossRef](#)] [[PubMed](#)]
161. Cui, K.; Zhao, K. Genome-wide approaches to determining nucleosome occupancy in metazoans using MNase-Seq. *Methods Mol. Biol.* **2012**, *833*, 413–419. [[PubMed](#)]
162. Song, L.; Crawford, G.E. DNase-seq: A high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. *Cold Spring Harb. Protoc.* **2010**, *2010*, 5381–5384. [[CrossRef](#)] [[PubMed](#)]
163. Dekker, J.; Rippe, K.; Dekker, M.; Kleckner, N. Capturing chromosome conformation. *Science* **2002**, *295*, 1306–1311. [[CrossRef](#)] [[PubMed](#)]
164. Loewen, G.; Zhuo, Y.; Zhuang, Y.; Jayawickramarajah, J.; Shan, B. lincRNA HOTAIR as a novel promoter of cancer progression. *J. Can. Res. Updates* **2014**, *3*, 134–140. [[PubMed](#)]
165. Hajjari, M.; Salavaty, A. HOTAIR: An oncogenic long non-coding RNA in different cancers. *Cancer Biol. Med.* **2015**, *12*, 1–9. [[PubMed](#)]
166. Gupta, R.A.; Shah, N.; Wang, K.C.; Kim, J.; Horlings, H.M.; Wong, D.J.; Tsai, M.C.; Hung, T.; Argani, P.; Rinn, J.L.; *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **2010**, *464*, 1071–1076. [[CrossRef](#)] [[PubMed](#)]
167. Ge, X.S.; Ma, H.J.; Zheng, X.H.; Ruan, H.L.; Liao, X.Y.; Xue, W.Q.; Chen, Y.B.; Zhang, Y.; Jia, W.H. HOTAIR, a prognostic factor in esophageal squamous cell carcinoma, inhibits WIF-1 expression and activates Wnt pathway. *Cancer Sci.* **2013**, *104*, 1675–1682. [[CrossRef](#)] [[PubMed](#)]
168. Padua Alves, C.; Fonseca, A.S.; Muys, B.R.; de Barros, E.L.B.R.; Burger, M.C.; de Souza, J.E.; Valente, V.; Zago, M.A.; Silva, W.A., Jr. Brief report: The lincRNA Hotair is required for epithelial-to-mesenchymal transition and stemness maintenance of cancer cell lines. *Stem Cells* **2013**, *31*, 2827–2832. [[CrossRef](#)] [[PubMed](#)]
169. Crea, F.; Clermont, P.L.; Parolia, A.; Wang, Y.; Helgason, C.D. The non-coding transcriptome as a dynamic regulator of cancer metastasis. *Cancer Metastasis Rev.* **2014**, *33*, 1–16. [[CrossRef](#)] [[PubMed](#)]
170. Wilkerson, M.D.; Cabanski, C.R.; Sun, W.; Hoadley, K.A.; Walter, V.; Mose, L.E.; Troester, M.A.; Hammerman, P.S.; Parker, J.S.; Perou, C.M.; *et al.* Integrated RNA and DNA sequencing improves mutation detection in low purity tumors. *Nucleic Acids Res.* **2014**, *42*, e107. [[CrossRef](#)] [[PubMed](#)]
171. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **2014**, *513*, 202–209.
172. Ali, S.M.; Sanford, E.M.; Klempner, S.J.; Rubinson, D.A.; Wang, K.; Palma, N.A.; Chmielecki, J.; Yelensky, R.; Palmer, G.A.; Morosini, D.; *et al.* Prospective comprehensive genomic profiling of advanced gastric carcinoma cases reveals frequent clinically relevant genomic alterations and new routes for targeted therapies. *Oncologist* **2015**, *20*, 499–507. [[CrossRef](#)] [[PubMed](#)]
173. Gajjar, A.; Pfister, S.M.; Taylor, M.D.; Gilbertson, R.J. Molecular insights into pediatric brain tumors have the potential to transform therapy. *Clin. Cancer Res.* **2014**, *20*, 5630–5640. [[CrossRef](#)] [[PubMed](#)]

174. Shah, S.P.; Morin, R.D.; Khattra, J.; Prentice, L.; Pugh, T.; Burleigh, A.; Delaney, A.; Gelmon, K.; Guliany, R.; Senz, J.; *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **2009**, *461*, 809–813. [[CrossRef](#)] [[PubMed](#)]
175. Davis, C.F.; Ricketts, C.J.; Wang, M.; Yang, L.; Cherniack, A.D.; Shen, H.; Buhay, C.; Kang, H.; Kim, S.C.; Fahey, C.C.; *et al.* The somatic genomic landscape of chromophobe renal cell carcinoma. *Cancer Cell* **2014**, *26*, 319–330. [[CrossRef](#)] [[PubMed](#)]
176. Gerlinger, M.; Rowan, A.J.; Horswell, S.; Larkin, J.; Endesfelder, D.; Gronroos, E.; Martinez, P.; Matthews, N.; Stewart, A.; Tarpey, P.; *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **2012**, *366*, 883–892. [[CrossRef](#)] [[PubMed](#)]
177. Wyatt, A.W.; Mo, F.; Wang, K.; McConeghy, B.; Brahmabhatt, S.; Jong, L.; Mitchell, D.M.; Johnston, R.L.; Haegert, A.; Li, E.; *et al.* Heterogeneity in the inter-tumor transcriptome of high risk prostate cancer. *Genome Biol.* **2014**, *15*, 421–426. [[CrossRef](#)] [[PubMed](#)]
178. Ding, L.; Ley, T.J.; Larson, D.E.; Miller, C.A.; Koboldt, D.C.; Welch, J.S.; Ritchey, J.K.; Young, M.A.; Lamprecht, T.; McLellan, M.D.; *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **2012**, *481*, 506–510. [[CrossRef](#)] [[PubMed](#)]
179. Macaulay, I.C.; Haerty, W.; Kumar, P.; Li, Y.I.; Hu, T.X.; Teng, M.J.; Goolam, M.; Saurat, N.; Coupland, P.; Shirley, L.M.; *et al.* G&T-seq: Parallel sequencing of single-cell genomes and transcriptomes. *Nat. Methods* **2015**, *12*, 519–522. [[PubMed](#)]
180. The Cancer Genome Atlas (TCGA): The next stage. Available online: http://cancergenome.nih.gov/newsevents/newsannouncements/TCGA_The_Next_Stage (accessed on 13 April 2015).
181. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **2008**, *455*, 1061–1068.
182. The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* **2011**, *474*, 609–615.
183. The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **2012**, *487*, 330–337.
184. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* **2012**, *489*, 519–525.
185. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **2012**, *490*, 61–70.
186. The Cancer Genome Atlas Research Network. Integrated genomic characterization of endometrial carcinoma. *Nature* **2013**, *497*, 67–73.
187. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **2013**, *368*, 2059–2074. [[CrossRef](#)] [[PubMed](#)]
188. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **2013**, *499*, 43–49.
189. The Cancer Genome Atlas Research Network. Comprehensive molecular characterization of urothelial bladder carcinoma. *Nature* **2014**, *507*, 315–322.
190. The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **2014**, *511*, 543–550.

191. The Cancer Genome Atlas Research Network. Integrated genomic characterization of papillary thyroid carcinoma. *Cell* **2014**, *159*, 676–690.
192. The Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **2015**, *517*, 576–582.
193. The Cancer Genome Atlas Research Network. Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas. *N. Engl. J. Med.* **2015**, *372*, 2481–2498. [[CrossRef](#)] [[PubMed](#)]
194. The Cancer Genome Atlas Network. Genomic Classification of Cutaneous Melanoma. *Cell* **2015**, *161*, 1681–1696.
195. The Cancer Genome Atlas Research Network; Weinstein, J.N.; Collisson, E.A.; Mills, G.B.; Shaw, K.R.M.; Ozenberger, B.A.; Ellrott, K.; Shmulevich, I.; Sander, C.; Stuart, J.M. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet.* **2013**, *45*, 1113–1120.
196. Hoadley, K.A.; Yau, C.; Wolf, D.M.; Cherniack, A.D.; Tamborero, D.; Ng, S.; Leiserson, M.D.M.; Niu, B.; McLellan, M.D.; Uzunangelov, V.; *et al.* Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* **2014**, *158*, 929–944. [[CrossRef](#)] [[PubMed](#)]
197. Liu, X.-S.; Haines, J.E.; Mehanna, E.K.; Genet, M.D.; Ben-Sahra, I.; Asara, J.M.; Manning, B.D.; Yuan, Z.-M. ZBTB7A acts as a tumor suppressor through the transcriptional repression of glycolysis. *Genes Dev.* **2014**, *28*, 1917–1928. [[CrossRef](#)] [[PubMed](#)]
198. Jamshidi, F.; Pleasance, E.; Li, Y.; Shen, Y.; Kasaian, K.; Corbett, R.; Eirew, P.; Lum, A.; Pandoh, P.; Zhao, Y.; *et al.* Diagnostic value of next-generation sequencing in an unusual sphenoid tumor. *Oncologist* **2014**, *19*, 623–630. [[CrossRef](#)] [[PubMed](#)]
199. Sheffield, B.S.; Tinker, A.V.; Shen, Y.; Hwang, H.; Li-Chang, H.H.; Pleasance, E.; Ch'ng, C.; Lum, A.; Lorette, J.; McConnell, Y.J.; *et al.* Personalized oncogenomics: Clinical experience with malignant peritoneal mesothelioma using whole genome sequencing. *PLoS ONE* **2015**, *10*, e0119689. [[CrossRef](#)] [[PubMed](#)]
200. Personalized Oncogenomics (POG) Program of British Columbia. Available online: <https://clinicaltrials.gov/ct2/show/NCT02155621> (accessed on 11 May 2015).
201. Ray, T. With new funds, BC Cancer Agency aims to sequence 300 patients for personalized oncogenomics trial. Available online: <https://www.genomeweb.com/genetic-research/new-funds-bc-cancer-agency-aims-sequence-300-patients-personalized-onco-genomics> (accessed on 11 May 2015).
202. Jones, S.J.; Laskin, J.; Li, Y.Y.; Griffith, O.L.; An, J.; Bilenky, M.; Butterfield, Y.S.; Cezard, T.; Chuah, E.; Corbett, R.; *et al.* Evolution of an adenocarcinoma in response to selection by targeted kinase inhibitors. *Genome Biol.* **2010**, *11*, 81–82. [[CrossRef](#)] [[PubMed](#)]
203. Welch, J.S.; Westervelt, P.; Ding, L.; Larson, D.E.; Klco, J.M.; Kulkarni, S.; Wallis, J.; Chen, K.; Payton, J.E.; Fulton, R.S.; *et al.* Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. *JAMA* **2011**, *305*, 1577–1584. [[CrossRef](#)] [[PubMed](#)]
204. Loaiza-Bonilla, A.; Clayton, E.; Furth, E.; O'Hara, M.; Morrisette, J. Dramatic response to dabrafenib and trametinib combination in a BRAF V600E-mutated cholangiocarcinoma: Implementation of a molecular tumour board and next-generation sequencing for personalized medicine. *Ecancermedicalscience* **2014**, *8*, 479. [[CrossRef](#)] [[PubMed](#)]

205. Sohn, I.; Jung, W.Y.; Sung, C.O. Somatic hypermutation and outcomes of platinum based chemotherapy in patients with high grade serous ovarian cancer. *Gynecol. Oncol.* **2012**, *126*, 103–108. [[CrossRef](#)] [[PubMed](#)]
206. Wagle, N.; Grabiner, B.C.; van Allen, E.M.; Hodis, E.; Jacobus, S.; Supko, J.G.; Stewart, M.; Choueiri, T.K.; Gandhi, L.; Cleary, J.M.; *et al.* Activating mTOR mutations in a patient with an extraordinary response on a phase I trial of everolimus and pazopanib. *Cancer Discov.* **2014**, *4*, 546–553. [[CrossRef](#)] [[PubMed](#)]
207. Iyer, G.; Hanrahan, A.J.; Milowsky, M.I.; Al-Ahmadie, H.; Scott, S.N.; Janakiraman, M.; Pirun, M.; Sander, C.; Socci, N.D.; Ostrovnya, I.; *et al.* Genome sequencing identifies a basis for everolimus sensitivity. *Science* **2012**, *338*, 221. [[CrossRef](#)] [[PubMed](#)]
208. Al-Ahmadie, H.; Iyer, G.; Hohl, M.; Asthana, S.; Inagaki, A.; Schultz, N.; Hanrahan, A.J.; Scott, S.N.; Brannon, A.R.; McDermott, G.C.; *et al.* Synthetic lethality in ATM-deficient RAD50-mutant tumors underlies outlier response to cancer therapy. *Cancer Discov.* **2014**, *4*, 1014–1021. [[CrossRef](#)] [[PubMed](#)]
209. Weber, A.M.; Ryan, A.J. ATM and ATR as therapeutic targets in cancer. *Pharmacol. Ther.* **2015**, *149*, 124–138. [[CrossRef](#)] [[PubMed](#)]
210. Kaelin, W.G., Jr. The concept of synthetic lethality in the context of anticancer therapy. *Nat. Rev. Cancer* **2005**, *5*, 689–698. [[CrossRef](#)] [[PubMed](#)]
211. Chan, S.M.; Thomas, D.; Corces-Zimmerman, M.R.; Xavy, S.; Rastogi, S.; Hong, W.J.; Zhao, F.; Medeiros, B.C.; Tyvoll, D.A.; Majeti, R. Isocitrate dehydrogenase 1 and 2 mutations induce BCL-2 dependence in acute myeloid leukemia. *Nat. Med.* **2015**, *21*, 178–184. [[CrossRef](#)] [[PubMed](#)]
212. Molenaar, R.J.; Radivoyevitch, T.; Maciejewski, J.P.; van Noorden, C.J.; Bleeker, F.E. The driver and passenger effects of isocitrate dehydrogenase 1 and 2 mutations in oncogenesis and survival prolongation. *Biochim. Biophys. Acta* **2014**, *1846*, 326–341. [[CrossRef](#)] [[PubMed](#)]
213. Brannon, A.R.; Sawyers, C.L. “N of 1” case reports in the era of whole-genome sequencing. *J. Clin. Invest.* **2013**, *123*, 4568–4570. [[CrossRef](#)] [[PubMed](#)]
214. Mullard, A. Learning from exceptional drug responders. *Nat. Rev. Drug Discov.* **2014**, *13*, 401–402. [[CrossRef](#)] [[PubMed](#)]
215. Takebe, N.; McShane, L.; Conley, B. Exceptional responders—discovering predictive biomarkers. *Nat. Rev. Clin. Oncol.* **2015**, *12*, 132–134. [[CrossRef](#)] [[PubMed](#)]
216. Molecular Case Studies. Available online: <http://molecularcasestudies.org/> (accessed on 18 August 2015).