Review

Analyzing Modern Biomolecules: The Revolution of Nucleic-Acid Sequencing – Review

Gabriel Dorado [1,*], Sergio Gálvez [2], Teresa E. Rosales [3], Víctor F. Vásquez [4] and Pilar Hernández [5]

[1]   Dep. Bioquímica y Biología Molecular, Campus Rabanales C6-1-E17, Campus de Excelencia Internacional Agroalimentario (ceiA3), Universidad de Córdoba, 14071 Córdoba, Spain
[2]   Dep. Lenguajes y Ciencias de la Computación, Boulevard Louis Pasteur 35, Universidad de Málaga, 29071 Málaga, Spain; galvez@uma.es
[3]   Laboratorio de Arqueobiología, Avda. Universitaria s/n, Universidad Nacional de Trujillo, 13011 Trujillo, Peru; teresa1905@hotmail.com
[4]   Centro de Investigaciones Arqueobiológicas y Paleoecológicas Andinas Arqueobios, Martínez de Companón 430-Bajo 100, Urbanización San Andres, 13088 Trujillo, Peru; vivasa2401@yahoo.com
[5]   Instituto de Agricultura Sostenible (IAS), Consejo Superior de Investigaciones Científicas (CSIC), Alameda del Obispo s/n, 14080 Córdoba, Spain; phernandez@ias.csic.es
[*]   Correspondence: bb1dopeg@uco.es; Tel.: +34-957218689

Abstract: Recent developments have revolutionized the study of biomolecules. Among them are molecular markers, amplification and sequencing of nucleic acids. The latter is classified into three generations. The first allows to sequence small DNA fragments. The second one increases throughput, reducing turnaround and pricing, and is therefore more convenient to sequence full genomes and transcriptomes. The third generation is currently pushing technology to its limits, being able to sequence single molecules, without previous amplification, which was previously impossible. Besides, this represents a new revolution, allowing researchers to directly sequence RNA without previous retrotranscription. These technologies are having a significant impact on different areas, such as medicine, agronomy, ecology and biotechnology. Additionally, the study of biomolecules is revealing interesting evolutionary information. That includes deciphering what makes us human, including phenomena like non-coding RNA expansion. All this is redefining the concept of gene and transcript. Basic analyses and applications are now facilitated with new genome editing tools, such as CRISPR. All these developments, in general, and nucleic-acid sequencing, in particular, are opening a new exciting era of biomolecule analyses and applications, including personalized medicine, and diagnosis and prevention of diseases for humans and other animals.

Keywords: first-generation sequencing (FGS); second-generation sequencing (SGS); third-generation sequencing (TGS); high-throughput sequencing (HTS); next-generation sequencing (NGS); structural genomics; functional genomics; epigenomics; metagenomics

1. Three Sequencing Generations

Analyses of biomolecules have been revolutionized by different technologies, including: (i) molecular-marker design; (ii) amplification of deoxyribonucleic acids (DNA); and (iii) nucleic-acid sequencing. The latter allows to read the code of life, being initially developed for DNA. That also allows to indirectly sequence ribonucleic acids (RNA), after retrotranscription into complementary DNA (cDNA). This is known by the misleading name of RNA sequencing (RNA-seq), instead of the more appropriate cDNA sequencing (cDNA-seq) terminology. Actually, it is not a true sequencing of native RNA, but of cDNA instead, with all biases that might be associated with such a process. Initially, all this required the previous amplification of DNA or cDNA by in vivo molecular cloning into suitable hosts, like *Escherichia coli*. Such processes typically required several years of dedicated work. The methodology was significantly enhanced by in vitro amplification

technologies, such as polymerase chain reaction (PCR). A significant step forward was accomplished with the development of platforms capable of massive parallel sequencing, as well as sequencing single molecules of nucleic acids. That way, it is now possible to directly sequence not only DNA, without previous amplification or labeling steps, but also RNA, without previous retrotranscription. Nucleic-acid sequencing technologies are classified as first-generation sequencing (FGS), second-generation sequencing (SGS) and third-generation sequencing (TGS). High-throughput sequencing (HTS) methodologies, such as SGS and TGS, are sometimes known with the ambiguous "next"-generation sequencing (NGS) terminology. Such platforms are briefly described below (Figure 1).

Nucleic acid

First-generation sequencing
Maxam-Gilbert Chemical-Degradation
Sanger Dideoxy-Terminator

Second-generation sequencing
Roche 454 Life Sciences Emulsion-PCR
Illumina Reversible-Terminator
Life Technologies SOLiD
Life Technologies Ion Torrent Chip

Third-generation sequencing
Helicos BioSciences True Single-Molecule
Pacific Biosciences Single-Molecule Real-Time
Complete Genomics Combinatorial Probe-Anchor Ligation
Oxford Nanopore Technologies

In vivo or *in vitro* amplification
(previous cDNA synthesis by retrotranscription in case of RNA)

Direct single-molecule sequencing
(true direct RNA sequencing)

Lowest throughput

Higher throughput

Highest throughput
Longer reads

Shorter reads

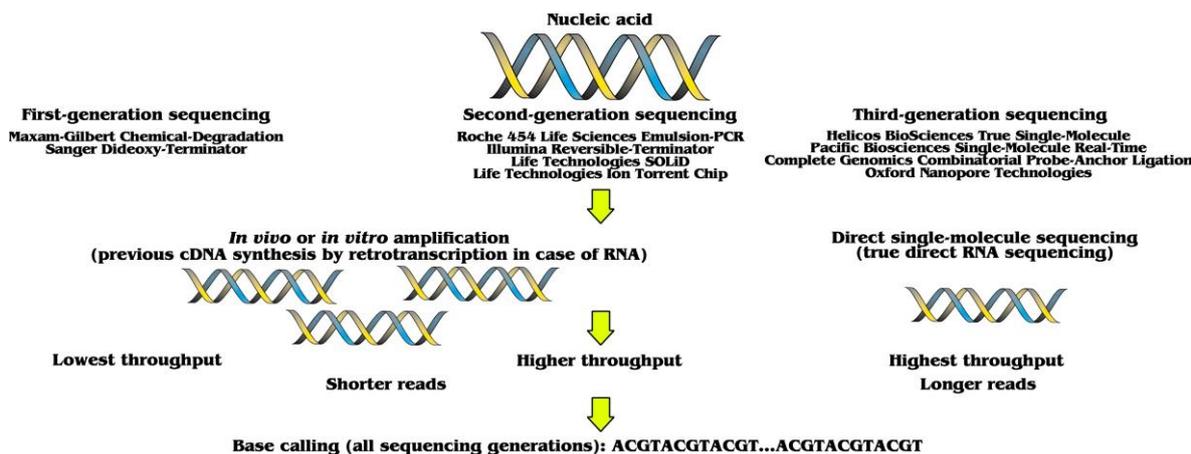**Base calling (all sequencing generations): ACGTACGTACGT...ACGTACGTACGT**

Figure 1. Comparison of sequencing platforms. FGS allows to sequence small fragments of DNA. SGS represents a significant increase in throughput. Finally, besides generating much longer reads, TGS can sequence single molecules without previous RNA retrotranscription or DNA amplification. Such a breakthrough allows to directly sequence RNA.

FGS platforms include (i) chemical degradation (CD; Maxam-Gilbert); and (ii) dideoxy terminator (ddT; Sanger). They can sequence short fragments of DNA. FGS methods were revolutionary when developed, since they allowed researchers to sequence DNA for the first time. Sanger's approach was further optimized (e.g., using fluorescent labels, instead of the original radioactive ones). In vitro amplification replaced tedious and time-consuming molecular cloning protocols, drastically reducing workflow times from several years to just months or minutes. Thus, it became very popular, being extensively used for decades to sequence short stretches of DNA. However, FGS approaches are expensive, time consuming and with low throughput. Therefore, they are not practical to sequence full genomes or transcriptomes. Indeed, the Human Genome Project using such a platform took 15 years, at a cost of three million milliard USD, even after optimizations that increased reading lengths and reduced errors, allowing researchers to finish it in half the time than previously expected at the time [1]. Bioinformatics tools were used to generate contigs, scaffolds, chromosome assemblies and full genome annotation, for such de novo sequencing. A large number of reactions and sequencing machines were used, as well as an intense labor force.

Subsequently, SGS of DNA represented a new breakthrough in biomolecule research, allowing to sequence genomes at an affordable time–cost scale. Indeed, SGS overcomes some limitations of FGS, using different approaches, corresponding to different commercial platforms, including: (i) emulsion PCR (emPCR; Roche-454 Life Sciences; Basel, Switzerland); (ii) reversible-terminator (RT; Illumina; San Diego, CA, USA); (iii) sequencing by oligonucleotide ligation and detection (SOLiD; Thermo Fisher Scientific-Life Technologies; Waltham, MA, USA); and (iv) ion torrent (IonT) chip, from the same manufacturer. Yet, albeit revolutionary in relation to FGS, SGS still has some shortcomings. They include the requirements to amplify DNA or retrotranscribe RNA. Indeed, that may introduce sequence biases, due to DNA polymerase or retrotranscriptase errors (generating mutations), with subsequent errors in the sequence readings [2]. Failure to properly read sequences

may also arise in repetitive stretches (including homopolymers) and CG-rich regions, due to enzymatic limitations of DNA polymerases. Besides, the typical short-readings of SGS may pose insurmountable hindrances, since they may be difficult, if not impossible, to be accurately assembled, mainly in the absence of a reference genome. The rationale is that similar or identical short fragments may be located at different genome sites. So, it may become impossible to map a particular short sequence to any specific site, amongst the multiple potential targets available in the genome [3]. Also, as with FGS, SGS can be applied to sequence DNA, but cannot directly sequence RNA molecules.

Fortunately, TGS of nucleic acids represents a new revolution [4]. Its key advantages stem from the fact that it can directly sequence long single nucleic-acid molecules. Thus, it allows true and direct RNA-sequencing (DRS) and direct DNA-sequencing (DDS) of molecules, without previous retrotranscription or amplification, respectively. Therefore, it prevents biases associated with such steps [2]. Several TGS platforms have been released, including: (i) true single-molecule sequencing (tSMS; Helicos BioSciences; Cambridge, MA, USA); (ii) single-molecule real-time (SMRT; Pacific Biosciences; PacBio; Menlo Park, CA, USA); (iii) combinatorial probe-anchor ligation (cPAL; BGI Group-Complete Genomics; Shenzhen, China); and (iv) nanopore (NP) sequencing (Oxford Nanopore Technologies; Oxford, UK). The approaches from Helicos and Oxford allow direct sequencing of DNA or RNA. Additionally, long-read sequencing platforms have great potential in many research areas [5,6], allowing annotations without, or with lower, assembly requirements (depending on the source sequence length), streamlining data processing workflows [7]. In particular, Pacific Biosciences generates long reads of 20 kb on average, reaching 300 kb [8]. Nanopore sequencing can generate 30 kb reads, reaching even 2.3 Mb [9]. However, some shortcomings of TGS (like the requirement for higher nucleic acid concentrations and higher error rates than other platforms) should be properly addressed, to reach its full potential [4,9–12].

2. Applications of Nucleic-Acid Sequencing

Optimizations in experimental protocols and improvement of commercial sequencing platforms have allowed an exponential growth of applications of nucleic-acid sequencing. Indeed, there is currently a new revolution, as shown by the exponential growth of publications, regarding the possibility to sequence DNA and RNA from since-cells, as well as single organelles (mitochondria and chloroplasts) [13]. Special emphasis is now focused on integrating different -omics technologies, such as genomics (usually, DNA), transcriptomics (RNA), proteomics (peptides, like proteins), epigenomics (epigenetic factors) and metabolomics (metabolites), that eventually influence phenotypes in health and disease [14–17]. Furthermore, a combination of multi-omics techniques, complemented with morphological and physiological ones, allows a holistic approach to deciphering biological systems [18,19].

The huge amount of data generated, mainly by SGS and TGS, is demanding new software and hardware developments. Thus, mathematical tools, including statistical and bioinformatics ones involving artificial intelligence (AI), machine learning (ML) and dedicated neural network hardware (like neural engines), are being developed to better analyze the big data generated [20,21]. Some bioinformatics tools have been designed to reduce sequencing errors, like the in vivo genome diversity analyzer (iGDA), which can identify low frequency (down to 0.2%) single-nucleotide polymorphisms (SNP) [12]. Besides, recent developments are allowing to enrich nucleic-acids from samples using genome-editing tools, like clustered regularly-interspaced short palindromic repeats (CRISPR) [22]. A recent example of the relevance of new nucleic-acid sequencing technologies can be illustrated with their use to fight the current pandemic of coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [23].

Interestingly, the National Aeronautics and Space Administration (NASA) <https://www.nasa.gov> (accessed on 27 July 2021) has recently tested the MinION Mk1B portable sequencer (handheld; dimensions of 10′5 × 3′3 × 2′3 cm and just 87 g of weight) from

Oxford Nanopore Technologies <https://nanoporetech.com/products/minion> (accessed on 27 July 2021) for astrobiology [24–27]. It can sequence nucleic acids in just 10 min, at an affordable price of just 1000 USD for the starter kit (including MinION and all materials for two runs). Traditionally, crew members of the International Space Station (ISS) have been routinely monitored for health status, including DNA tests. This requires sending samples to planet Earth for analyses. Since the MinION works in microgravity, it allows the identification of biological entities and the diagnosis of diseases in space. It could be also used in future missions to Mars or other places, allowing to search for and identify nucleic-acid-based life on such places [24]. Of course, these are uncertain astrobiology projects. Indeed, if it exists, finding life outside our planet is not an easy task. Time will tell, but such a miniature sequencer also has interesting applications on Earth, including in situ ecological studies. Some significant applications of nucleic-acid sequencing are described below.

2.1. Structural Genomics

Nucleic-acid sequencing allows the identification of specific nucleotide sequences of biological entities. That is interesting per se, as well as to compare mutations (polymorphisms) between molecules (genotyping). There is a plethora of applications of structural genomics, including, among others: (i) comparative genomics, to discover identities and differences between molecules; (ii) chromatin profiling, to identify regulatory regions; (iii) diagnostic and treatment of diseases, with great potential for agronomy, pharmacology and medicine; (iv) marker-assisted breeding, significantly accelerating selection; (v) certification of protected designations of origin (PDO), protected geographical indication (PGI) and traditional specialties guaranteed (TSG) for foodstuffs; (vi) identification of contaminations and frauds in foodstuffs; (vii) illegal traffic monitoring, e.g., protected species and their remains; (viii) biodiversity and ecological research, including management of germplasm banks; (ix) linking genotypes to phenotypes, including behavior; (x) bioengineering, with great impact on agronomy, medicine and biotechnology; (xi) origin of life studies; and (xii) synthetic biology, further allowing the investigation of the origin of life, and also with significant biotechnological potential. Nucleic-acid sequencing is relevant when studying any biological entity or its parts, virtually covering all life-science-related areas. To illustrate such applications, some examples of this revolution in biomolecule analyses are described below, with emphasis on the most recent ones, mostly related to medical applications.

As an example of the relevance of structural genomics, the Human Genome Project opened the door for whole-genome resequencing and targeted applications, such as exome resequencing. This has important implications in disease diagnostics and clinical treatments. Its full potential is being currently expanded with SGS and TGS platforms. This should allow further accomplishments, with the promise of 100 USD human genome resequencing. Genotyping is traditionally carried out using molecular markers or sequencing specific targeted common/known loci. Whole-genome sequencing (WGS) represents the ultimate molecular marker, allowing such genetic profiling with an unprecedented power. This includes different biotechnological areas, such as pharmacogenetic profiling [28]. Indeed, twins and even two cells from the same organism can now be differentiated with such a powerful tool. In this manner, new sequencing technologies are allowing researchers to better diagnose and analyze diseases [29]. Amongst the many examples available are the fight against complex diseases such as cancer [13,30] and neuromuscular disorders (NMD), involving more than 600 genes, affecting one in every thousand persons worldwide [31], and structural variations (SV), as shown for conditions such as autism. Interestingly, some of them are related to non-coding sequences [32].

Besides nuclear DNA in eukaryotes, organelle genomes should also be considered. For instance, they are relevant when analyzing mitochondrial disorders. New sequencing platforms have revolutionized diagnostics of such diseases, mainly exome and whole-genome approaches, including mitochondrial heteroplasmy [33]. Nevertheless, a holistic

-omics approach is needed to generate more comprehensive results, also requiring new bioinformatics tools to properly analyze them [34–38].

New sequencing technologies are also allowing to study beneficial and pathogenic biological entities, representing significant advances for medical diagnosis and therapy [39], as well as agronomy [40,41], allowing researchers to sequence even single cells [42]. Horizontal gene transfer (HGT) in microbial communities is also important. This can generate antibiotic resistance, with significant relevance in different research areas [43]. Additionally, another of the most interesting applications of genome sequencing is personalized medicine, like sequencing single gametes [44,45]. Nucleic acids can also be used to store any kind of information in a compact and efficient way which can be retrieved by sequencing and decoding [46].

2.2. Functional Genomics

Transcriptomics was initially addressed retrotranscribing RNA into cDNA and further in vivo molecular cloning. That allowed the sequencing of specific molecules using FGS. The procedure was significantly optimized with in vitro amplification methodologies, such as PCR. Furthermore, SGS opened the door to sequencing full transcriptomes at an affordable cost, which was another revolution in biomolecule research. However, the most significant breakthrough came from TGS, since it allowed the direct sequencing of RNA, without retrotranscription or amplification steps, avoiding the biases related to them. Like structural genomics described above, functional genomics or transcriptomics are used in different fields, such as agronomy and medicine. Abiotic and biotic stresses, as well as disease tolerance and resistance, can be analyzed in plants and animals at the molecular level, with significant implications in breeding programs and health [47]. Such strategies can be coupled with ML to optimize big data analyses [48,49]. Genomics-assisted breeding (GAB) allows to improve the germplasm [50]. Besides, multiple stress combinations can be studied [51]. Systems biology strategies are particularly interesting, implementing holistic approaches in these scenarios, integrating different -omics and bioinformatics tools [52]. This is especially relevant in the current trend of global warming and climate change [53–57]. As with structural genomics, studies of functional genomics are growing at an exponential rate in different areas related to biological entities. Some relevant examples are described below, with emphasis on medical applications.

New sequencing platforms, in general, and TGS, in particular, with longer reads of full-length transcripts, are revealing new genes [58]. Bioinformatics tools have been developed to correct errors for such platforms [58], allowing reference-free transcriptome analyses [6,59]. This is particularly useful when studying RNA isoforms generated by alternative splicing (AS). Its dysregulation may be responsible for initiation and progression of diseases like cancer. Thus, specific computational tools have been developed to integrate genomics and transcriptomics, for a proper characterization of alternative splicing in health and disease [60], including mitochondrial diseases [34,61]. In relation to that, long-read isoform quantification and analysis (LIQA) allows to identify differential alternative splicing (DAS). Such tools have been applied to study splicing events in cancer [62]. ML approaches, such as deep learning (DL), have been used to analyze the effect of disrupting splicing on pathogenicity [63]. New sequencing technologies also allow novel immunotherapy strategies, to fight cancer and other complex diseases. Interestingly, cancer cells usually exhibit transcriptomics dysregulation. In this scenario, tumor antigens (TA) can be designed from aberrant transcripts encoding cancer-specific proteins. Additionally, big data approaches are used to analyze multi-omics data from cancer cells. Such knowledge allows translating experimental results into new, more efficient therapies with an unprecedented power [64].

Total RNA, poly(A) RNA and non-coding RNA populations can be isolated from tissues or cell cultures. Yet, such approaches can only generate average results, corresponding to such cell populations. Fortunately, it is now possible to isolate RNA from single cells and even from single nuclei. That allows an unprecedented dissection of transcription within millions of individual cells. Both single-cell RNA sequencing (scRNA-seq)

and single-nucleus RNA sequencing (snRNA-seq) have exciting applications [65–77], for instance: (i) discovering and characterizing cell type in health and diseases, such as cancer [13,78–89], with implications in immunology [90], immune-mediated diseases [91], immunotherapy [92–100] and drug resistance [101]; (ii) deciphering the roles of such specific cell types in health and disease [102], including mitochondrial heteroplasmy [33]; and (iii) analyzing cell emergence, development and plasticity in tissues and organisms. These studies are also applied to study plant biology [103,104]. Currently, sc/snRNA-seq are extensively being used in neuroscience research, including analyses of neurodegenerative disorders at the molecular level. This includes Parkinson's disease (PD) [105] and Alzheimer's disease (AD) [106]. Likewise, the development of the human brain from fetal to adult stages has been analyzed at the single-cell level. Interestingly, spatial transcriptomics allows to generate location maps of gene expression within cells, tissues, organs and whole organisms, comparing health and disease [66]. This can be done using probes with single-molecule fluorescence in situ hybridization (smFISH) [107], as well as sequencing with Slide-seq, which has ~10 μm spatial resolution [108,109].

On the other hand, cell identity is determined in different ways, with transcription factor (TF) networks playing an essential role. Recent developments in nucleic-acid sequencing, in general, and sc/snRNA-seq, in particular, allow to couple transcriptomic maps with cell identity, defining profiles of gene expression for each cell [110–115]. Interestingly, although pseudogenes were considered functionless, TGS has allowed to identify many transcribed pseudogenes, including protein-coding ones in normal and cancer human cells [116]. Transcriptomics has also been used to study cellular communications, including both intra- and inter-cellular signaling networks [117–122]. On the other hand, genome-editing technologies such as CRISPR can be combined with scRNA-seq applied to animal models and human organoids, to shed light on poorly understood diseases like autism [123]. Interestingly, non-coding sequences may be linked to some diseases [32]. As with structural genomics, organelle transcriptomics and mitochondrial disorders are also related to non-coding RNA [37]. Recently, TGS has allowed the sequencing of a class of them known as circular RNA (circRNA), which was previously refractory to sequencing [124].

It should also be taken into account that different sequencing platforms have advantages and disadvantages. Therefore, a combination of several of them may be needed for a comprehensive analysis of gene expression [125]. Besides, computational models [126], such as ML, have been applied to these studies [127], including dimension reduction methods [128]. Bioinformatics developments have also allowed to deconvult heterogeneous cell samples [129], as well as identify pathways or biological processes from transcriptomics [130]. As an example, the worldwide impact of rare diseases is significant, affecting ~350 million people. Nearly 6000 of them have been characterized at the molecular level, but diagnosis remains challenging. Thanks to the new sequencing developments, transcriptomics coupled with ML are being used to diagnose diseases, in general, and rare disorders, in particular [131].

2.3. Epigenomics

Epigenetic modifications may change chromosomal architectures, without modifying nucleic acid sequences. Depending on the cell type (prokaryote or eukaryote), different mechanisms may be involved in epigenetics, such as DNA methylation and histone acetylation, modulating different activities. Prokaryotic chromosomes lack histones. Therefore, DNA methylation is a main epigenetic regulator in such cells. There are three types of DNA methylation in prokaryotes: 6-methyladenine (6 mA), 4-methylcytosine (4 mC) and 5-methylcytosine (5 mC), including both bacteria and archaea. New sequencing technologies have allowed to characterize prokaryotic epigenomes [132], with recent developments such as Nick-seq. Thus, datasets are mined to increase sensitivity, specificity and accuracy. This way, genomic maps of DNA modifications and damage are generated, with single-nucleotide resolution [133]. Other new technology allows identification of sulfur re-

placing nonbridging phosphate oxygen, which is common in prokaryotes, through selective fluorescent labeling of single-stranded DNA phosphorothioate (PT) modifications [134].

The development of TGS capable of reading single molecules has allowed a comprehensive study of frequency and distribution of epigenetic modifications. This way, it has been possible to discover that they may be related to different functions, including regulation of gene expression, maintenance of genome stability, cell cycle, sporulation, cell shape, biofilm formation, motility, siderophore generation, membrane vesicle production, defense (discriminating self from non-self DNA, like the bacteriophages that can be cut by restrictases), lysogenicity, virulence (including pathogen–host interactions and host colonization) and response to the environment [132,135–139]. These studies are important to identify beneficial, harmless, opportunistic and pathogenic-virulent phenotypes related to health and disease [140]. For instance, it has been proposed that epigenetics are involved in the health effects of probiotics [141]. On the other hand, the relevance of DNA methylation in microorganism toxicity has been demonstrated in relation to *Escherichia coli* strains producing Shiga toxin. Indeed, they were responsible for ice cream- and lettuce-associated outbreaks in Belgium and the USA, respectively [142].

Likewise, it has been shown that inactivating 4 mC methyltransferase in *Leptospira* spp. pathogens produced genome-wide dysregulation of gene expression. Epigenetic studies have been also carried out with *Mycobacterium tuberculosis*, which is the infectious agent causing tuberculosis [143]. These findings are particularly relevant in the current trend of antibiotic resistance, with increasing numbers of total drug-resistant (TDR) bacteria resistant to all known antibiotics (known as "super bugs") [144]. Indeed, TDR *Mycobacterium tuberculosis* strains have arisen in the last two decades, mainly due to the misuse and abuse of antibiotics. This highlights the need for new prevention and treatment strategies for pathogenic bacteria, finding alternatives to antibiotics. The new sequencing technologies are being used to reach such a goal [145]. In this scenario, highly conserved DNA methyltransferases (MTases) are potential targets for epigenetic inhibitors to fight infections [139]. Besides, they may have potential biotechnological applications [146]. Additionally, they represent a valuable tool for aligning metagenomic contigs and scaffolds, preventing errors, as well as assigning mobile genetic elements (MGE), such as transposable elements (TE), to their host genomes [135].

Epigenetics is also important in plants. Being sessile organisms, they have developed regulatory mechanisms to fight abiotic and biotic stresses. This way, approaches such as DeMEter (DME) coupled with quantitative PCR (DME-qPCR) have been developed to quantify DNA methylation in plants. This has been demonstrated in Arabidopsis (*Arabidopsis thaliana*) and tomato (*Solanum lycopersicum*) [147]. On the other hand, 5 mC is involved in regulation of gene expression, repair, replication, transcription, recombination and transposon suppression in plants. The new sequencing platforms have allowed researchers to discover that 6 mA upregulates gene expression, both in eudicots, such as Arabidopsis, as well as monocots, such as rice (*Oryza sativa*) [21]. On the other hand, transposable elements may allow selective advantages and evolution in plants. However, they can also be harmful to their genome integrity, if not properly controlled. The latter can be accomplished through DNA methylation. Thus, it has been found that both 6 mA and 4 mC are involved in TE control of fig tree (*Ficus carica*) [148]. Interestingly, some stress responses are memorized (somatic epigenetic memory), and sometimes they are even inherited through meiosis (transgenerational epigenetic inheritance). This has potential applications to engineer stress-tolerant crops, especially in the current trend of global warming and climate change [149].

Virulence, as well as host and environmental adaptation of different plant pathogens, is also modulated by epigenetics. Examples include fungi and fungi-like microorganisms, such as *Phytophthora* spp. [150]. Interestingly, epigenetics can also be used to protect crops, using sustainable and ecologically-safe biocontrol strategies. For instance, TGS has been used to study biopesticides based on plant growth-promoting rhizobacteria (PGPR) such as *Bacillus velezensis* [151].

Additionally, epigenetics is directly and indirectly related to evolution, enhancing phenotypic plasticity [152], such as thermal adaptation. In this scenario, it is especially relevant for adaptation to present and future environmental conditions [153]. New sequencing methodologies allow to study epigenomics with an unprecedented resolutive power, including reduced-representation bisulfite sequencing (RRBS) and whole-genome bisulfite sequencing (WGBS), analyzing full genomes [154]. This has significant implications in many areas, such as ecology [155], environmental pollution including radiation [156,157], with relevant implications for cancer radioresistance [158] and health [135,138], as well neuropsychiatric disorders [159]. Besides, it has been found that mechanotransduction is involved in mechanical regulation of transcription and the epigenome, having a key role in cancer progression [160]. Interestingly, there is also a link between DNA damage and epigenetics. In this way, it has been found that 8-oxo-7,8-dihydro-2′-deoxyguanosine (8-oxodG) may modulate epigenetic regulation of gene expression [161].

Besides, as with genomics and transcriptomics, mitochondrial diseases have also been linked to organelle epigenetics [37]. Likewise, it is possible to study epigenomes of organisms, tissues, cells and cellular compartments and organelles such as nuclei, mitochondria and chloroplasts. Indeed, whole genome bisulfite sequencing has allowed researchers to demonstrate that methylation patterns are cell type-specific [162]. That opens the door to decipher how genomic regulatory networks work [102]. Interestingly, these findings are particularly relevant for personalized treatments of complex diseases, such as cancer, diabetes and asthma, as well as chronic age-related diseases, due to the interaction of multiple genetic and environmental factors [13,163,164]. Indeed, new sequencing technologies have allowed epigenetic profiling of different cancers [78,165]. It has been proposed that DNA methylation of probiotics plays an important role in immune responses of allergies, autoimmune disorders and cancer. This is mediated by regulatory T cells (Tregs). They are responsible for maintaining tolerance to self-antigens, preventing autoimmune diseases [166]. Treg cells are also subjected to epigenetic regulation. Therefore, an appropriate regulation in such cells, gut microbiota and their interaction is of paramount importance to maintain Treg function, preventing diseases. This is accomplished through transcriptional and epigenetic regulation [167].

On the other hand, developmental trajectories have also been studied. In this manner, it has been possible to identify particular cells responsible for expressing genes related to neurodevelopmental diseases [168], as well as changes during learning and memory [169]. Also, epigenetics have been related to dementia, such as Alzheimer's disease [170]. Such epigenetic modifications can be quantified not only in the central nervous system (CNS), but also in the cerebrospinal fluid. That opens the door for the development of biomarkers for early detection and treatment of AD [171]. Nevertheless, new bioinformatics developments are still needed, integrating multiplexed assays to better analyze health and disease [19]. An example in such direction is GermLine cycle Expression Analysis and Epigenetics (GLEANE) [172].

Epigenetics can also be applied to study environmental genotoxins causing mutations and cancer. Among them is acrylamide, which can be generated in foodstuff and beverages subjected to high temperatures, as happens with fried potatoes or coffee [173,174]. Acrylamide may generate brain tumors in general and glioblastoma in particular. This is the most aggressive and invasive brain tumor, with a life expectancy between one and one and a half years. Fortunately, new sequencing platforms such as SGS and TGS are significantly increasing our understanding of such diseases. This allows designing molecular markers and analyzing epigenetic profiling at the single-cell level, for better diagnostics, prevention and treatment [20]. On the other hand, recent discoveries have shown that epigenomics, in general, and social epigenomics, in particular, can also be used to ascertain how adverse social factors can generate diseases, especially in childhood [175]. Computational, statistical and bioinformatics tools are also needed to fully analyze epigenetics. In this scenario, as reported for transcriptomics, epigenetics has also been linked to rare diseases using ML, and particularly DL, approaches [176].

2.4. Metagenomics

Microbial communities are relevant in different areas, including human and animal medicine, food technology, agronomy, aquaculture and ecology. This way, they have important implications in health and disease, optimizing food and foodstuff production, breeding, biodiversity protection and the fight against the current trend of global warming and climate change. The new sequencing methodologies are opening the door to an unprecedented, powerful study of microbial communities [177,178]. In this way, many new species have been discovered [179,180]. This is contributing to identify healthy microbiomes, as well as diseases linked to dysbiosis scenarios [181]. Altered microbiome profiles have been found in many diseases, not only for typical infections, but also for other disfunctions, such as cancer [165,182]. Nevertheless, results obtained in different experiments may be different, due to experimental biases that must be properly addressed [183]. As with other genomic, transcriptomic and epigenomic areas, microbiome analyses (microbiomics) require appropriate bioinformatics tools [184]. TGS is particularly useful in metagenomic analyses, since it can be used to generate almost or even complete genomes with single reads, significantly reducing or not requiring contig assembly [185]. Therefore, TGS platforms are being used to find microorganisms present in human microbiomes, foodstuff and beverages like milk, aquaculture, soil and many other ecological niches, allowing to identify both beneficial and pathogenic microorganisms [186–188], including serotypes with closely related, or even the same, antigenic formulae [189].

Additionally, metagenomics can be used to study biological entities like virusoids, viroids plasmids and viruses [190], including viral quasispecies [191]. For instance, viruses responsible for hepatitis have been identified with short-read sequencing [192]. Long-read sequencing is even better, allowing single reads of full genomes. However, they may require high DNA concentrations, generating more sequencing errors than short-read platforms. Specific workflows combining wet-lab and bioinformatics pipelines have been developed to overcome these limitations. An example of such a strategy is viral metagenomics via MinION sequencing 2 (VirION2). Likewise, bioinformatics tools have been developed to increase long-read quality of sequencing [193]. As expected, short-read sequencing approaches failed to identify biodiversity that was found by long-read platforms, showing significantly higher biodiversity. The methodology has been further optimized to use samples with low nucleic acid concentrations, which may be especially relevant for environmental studies [194].

As with other biological systems, multiple-omics technologies open the door to longitudinal holistic approaches of microbial genomics. Thus, metagenomics, metatranscriptomics, metaproteomics and meta-metabolomics allow to generate an integrated picture of structure, function and phenotype. This opens the door to identify new functions, and even previously unknown species, with a better understanding and prediction of microbe–microbe and microbe–host interactions, with important microbiological, medical, agronomical and biotechnological implications [67,195].

3. Future Prospects and Concluding Remarks

The future is certainly promising for nucleic-acid sequencing, mostly due to the ingenious developments of new technologies. One interesting application area of nucleic-acid sequencing is food biotechnology, to identify pathogens. As an example, the IBM DNA Transistor <https://www.ibm.com/ibm/history/ibm100/us/en/icons/dnatransistor> (accessed on 27 July 2021), is being co-developed with Roche to identify pathogens in milk, as well as early detection, prevention, and personalized treatment of diseases. As Gustavo Stolovitzky (Manager of Functional Genomics and Systems Biology Group at IBM) said: "What is the next big thing in biotechnology? The answer is kind of simple if you're in the field—you need to know how to sequence DNA, fast and cheap". On the other hand, since TGS allows researchers to directly sequence single molecules, without biases associated with retrotranscription and amplification, that opens new fields of functional genomics. All these breakthroughs, coupled with fewer starting materials required, longer reads and

faster turnaround at lower prices, should boost scientific research and discoveries in areas related to living entities. These include medicine, agronomy, ecology and biotechnology.

These developments are relevant, not just for single specimens, but also for population studies, from microbes (metagenomics) to other analyses involving plants and animals. Technological developments and optimizations should generate more detailed and accurate results, allowing researchers to reach new insights and draw more accurate conclusions. In this manner, previously unattainable projects may be possible, for instance, to directly sequence nucleic acids when they are so scarce that FGS and SGS may generate negative results, since TGS can sequence single molecules. Likewise, deciphering what made us human is a provocative topic in biomolecular research, among other exciting research goals, in relation to the new sequencing platforms. In particular, research on non-coding RNA (which typically are short molecules) is particularly exciting, given the surprising implications of spurious or pervasive transcription in organic and cognitive evolution [196,197]. In this way, recent discoveries accomplished by nucleic-acid sequencing are redefining the concepts of gene and transcript.

All these developments, in general, and nucleic-acid sequencing, in particular, coupled with genome-editing breakthroughs, such as CRISPR, are highlighting the relevance of biomolecule analyses and applications. One of the goals is to re-sequence the human genome from the current 1000 USD price to just 100 USD, as shown by the National Human Genome Research Institute (NHGRI): "The Cost of Sequencing a Human Genome" <https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost> (accessed on 27 July 2021) [198]. Thus, everyone could have their genome sequenced in the near future. The implications for truly personalized medicine, with much more accurate and efficient diagnosis, prevention and treatment of diseases, will be unprecedented. This includes humans and other animals (veterinary medicine).

Additionally, associating nucleic-acid sequencing to activity-dependent labeling should allow to link transcriptomics and epigenomics with important functional implications, including roles of cells in physiology. New insights will be reached unifying nucleic-acid sequencing with functional, physiological, morphological and phenotypic data. All such research is now generating and will continue to produce huge amounts of data, requiring new software and hardware developments to properly analyze them. This includes AI, ML, DL and neural network chips, such as neural engines. Furthermore, new frameworks will be required to systematically filter, sort and organize such vast knowledge. This should make it easily available in a graphical way, for easier visualization and interpretation. It is clear now that this century will be revolutionary for several scientific areas, including molecular biology and biotechnology related to biomolecule research, with important implications and applications.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of Interest.

References

1. Lario, A.; Gonzalez, A.; Dorado, G. Automated laser-induced fluorescence DNA sequencing: Equalizing signal-to-noise ratios significantly enhances overall performance. *Anal. Biochem.* **1997**, *247*, 30–33. [CrossRef]
2. Ozsolak, F.; Platt, A.R.; Jones, D.R.; Reifenberger, J.G.; Sass, L.E.; McInerney, P.; Thompson, J.F.; Bowers, J.; Jarosz, M.; Milos, P.M. Direct RNA sequencing. *Nature* **2009**, *461*, 814–818. [CrossRef] [PubMed]

3. Heydari, M.; Miclotte, G.; Van de Peer, Y.; Fostier, J. Illumina error correction near highly repetitive DNA regions improves de novo genome assembly. *BMC Bioinform.* **2019**, *20*, 1–13. [CrossRef]

4. Bleidorn, C. Third generation sequencing: Technology and its potential impact on evolutionary biodiversity research. *Syst. Biodivers.* **2016**, *14*, 1–8. [CrossRef]

5. Blom, M.P.K. Opportunities and challenges for high-quality biodiversity tissue archives in the age of long-read sequencing. *Mol. Ecol.* **2021**. [CrossRef] [PubMed]

6. Broseus, L.; Thomas, A.; Oldfield, A.J.; Severac, D.; Dubois, E.; Ritchie, W. TALC: Transcript-level Aware Long-read Correction. *Bioinformatics* **2020**, *36*, 5000–5006. [CrossRef] [PubMed]

7. Du, N.; Shang, J.Y.; Sun, Y.N. Improving protein domain classification for third-generation sequencing reads using deep learning. *Bmc Genom.* **2021**, *22*, 1–13. [CrossRef] [PubMed]

8. Hestand, M.S.; Ameur, A. The Versatility of SMRT Sequencing. *Genes* **2019**, *10*, 24. [CrossRef]

9. Amarasinghe, S.L.; Su, S.; Dong, X.Y.; Zappia, L.; Ritchie, M.E.; Gouil, Q. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* **2020**, *21*, 1–16. [CrossRef] [PubMed]

10. Jain, M.; Koren, S.; Miga, K.H.; Quick, J.; Rand, A.C.; Sasani, T.A.; Tyson, J.R.; Beggs, A.D.; Dilthey, A.T.; Fiddes, I.T.; et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat. Biotechnol.* **2018**, *36*, 338–345. [CrossRef] [PubMed]

11. Wang, L.T.; Qu, L.; Yang, L.S.; Wang, Y.Y.; Zhu, H.Q. NanoReviser: An Error-Correction Tool for Nanopore Sequencing Based on a Deep Learning Algorithm. *Front. Genet.* **2020**, *11*, 900. [CrossRef]

12. Feng, Z.X.; Clemente, J.C.; Wong, B.; Schadt, E.E. Detecting and phasing minor single-nucleotide variants from long-read sequencing data. *Nat. Commun.* **2021**, *12*, 1–13. [CrossRef]

13. Bai, X.; Li, Y.X.; Zeng, X.M.; Zhao, Q.; Zhang, Z.W. Single-cell sequencing technology in tumor research. *Clin. Chim. Acta* **2021**, *518*, 101–109. [CrossRef] [PubMed]

14. Graw, S.; Chappell, K.; Washam, C.L.; Gies, A.; Bird, J.; Robeson, M.S.; Byrum, S.D. Multi-omics data integration considerations and study design for biological systems and disease. *Mol. Omics.* **2021**, *17*, 170–185. [CrossRef]

15. Reiter, T.; Brooks, P.T.; Irber, L.; Joslin, S.E.K.; Reid, C.M.; Scott, C.; Brown, C.T.; Pierce-Ward, N.T. Streamlining data-intensive biology with workflow systems. *Gigascience* **2021**, *10*, giaa140. [CrossRef]

16. Li, Y.; Ma, A.J.; Mathe, E.A.; Li, L.; Liu, B.Q.; Ma, Q. Elucidation of Biological Networks across Complex Diseases Using Single-Cell Omics. *Trends Genet.* **2020**, *36*, 951–966. [CrossRef] [PubMed]

17. Khella, C.A.; Mehta, G.A.; Mehta, R.N.; Gatza, M.L. Recent Advances in Integrative Multi-Omics Research in Breast and Ovarian Cancer. *J. Pers. Med.* **2021**, *11*, 149. [CrossRef]

18. Zhu, C.X.; Preissl, S.; Ren, B. Single-cell multimodal omics: The power of many. *Nat. Methods* **2020**, *17*, 11–14. [CrossRef] [PubMed]

19. Philpott, M.; Cribbs, A.P.; Brown, T.; Brown, T.; Oppermann, U. Advances and challenges in epigenomic single-cell sequencing applications. *Curr. Opin. Chem. Biol.* **2020**, *57*, 17–26. [CrossRef] [PubMed]

20. Jovcevska, I. Next Generation Sequencing and Machine Learning Technologies Are Painting the Epigenetic Portrait of Glioblastoma. *Front. Oncol.* **2020**, *10*, 798. [CrossRef]

21. Chachar, S.; Liu, J.R.; Zhang, P.X.; Riaz, A.; Guan, C.F.; Liu, S.Y. Harnessing Current Knowledge of DNA N6-Methyladenosine From Model Plants for Non-model Crops. *Front. Genet.* **2021**, *12*, 668317. [CrossRef] [PubMed]

22. Schultzhaus, Z.; Wang, Z.; Stenger, D. CRISPR-based enrichment strategies for targeted sequencing. *Biotechnol. Adv.* **2021**, *46*, 107672. [CrossRef]

23. Chiara, M.; D'Erchia, A.M.; Gissi, C.; Manzari, C.; Parisi, A.; Resta, N.; Zambelli, F.; Picardi, E.; Pavesi, G.; Horner, D.S.; et al. Next generation sequencing of SARS-CoV-2 genomes: Challenges, applications and opportunities. *Brief. Bioinform.* **2021**, *22*, 616–630. [CrossRef] [PubMed]

24. Castro-Wallace, S.L.; Chiu, C.Y.; John, K.K.; Stahl, S.E.; Rubins, K.H.; McIntyre, A.B.R.; Dworkin, J.P.; Lupisella, M.L.; Smith, D.J.; Botkin, D.J.; et al. Nanopore DNA Sequencing and Genome Assembly on the International Space Station. *Sci. Rep.* **2017**, *7*, 1–12. [CrossRef] [PubMed]

25. John, K.K.; Botkin, D.S.; Burton, A.S.; Castro-Wallace, S.L.; Chaput, J.D.; Dworkin, J.P.; Lehman, N.; Lupisella, M.L.; Mason, C.E.; Smith, D.J.; et al. The Biomolecule Sequencer Project: Nanopore sequencing as a dual-use tool for crew health and astrobiology investigations. In Proceedings of the 47th Lunar and Planetary Science Conference, The Woodlands, TX, USA, 21–25 March 2016.

26. Wong, S. Diagnostics in space: Will zero gravity add weight to new advances? *Expert Rev. Mol. Diagn.* **2020**, *20*, 1–4. [CrossRef]

27. Stahl-Rommel, S.; Jain, M.; Nguyen, H.N.; Arnold, R.R.; Aunon-Chancellor, S.M.; Sharp, G.M.; Castro, C.L.; John, K.K.; Juul, S.; Turner, D.J.; et al. Real-Time Culture-Independent Microbial Profiling Onboard the International Space Station Using Nanopore Sequencing. *Genes* **2021**, *12*, 106. [CrossRef]

28. Caspar, S.M.; Schneider, T.; Stoll, P.; Meienberg, J.; Matyas, G. Potential of whole-genome sequencing-based pharmacogenetic profiling. *Pharmacogenomics* **2021**, *22*, 177–190. [CrossRef]

29. Gorcenco, S.; Ilinca, A.; Almasoudi, W.; Kafantari, E.; Lindgren, A.G.; Puschmann, A. New generation genetic testing entering the clinic. *Parkinsonism Relat. D* **2020**, *73*, 72–84. [CrossRef]

30. Duan, Q.K.; Tang, C.; Ma, Z.; Chen, C.G.; Shang, X.B.; Yue, J.; Jiang, H.J.; Gao, Y.; Xu, B. Genomic Heterogeneity and Clonal Evolution in Gastroesophageal Junction Cancer Revealed by Single Cell DNA Sequencing. *Front. Oncol.* **2021**, *11*, 1574. [CrossRef] [PubMed]

31. Barp, A.; Mosca, L.; Sansone, V.A. Facilitations and Hurdles of Genetic Testing in Neuromuscular Disorders. *Diagnostics* **2021**, *11*, 701. [CrossRef]

32. Begum, G.; Albanna, A.; Bankapur, A.; Nassir, N.; Tambi, R.; Berdiev, B.K.; Akter, H.; Karuvantevida, N.; Kellam, B.; Alhashmi, D.; et al. Long-Read Sequencing Improves the Detection of Structural Variations Impacting Complex Non-Coding Elements of the Genome. *Int. J. Mol. Sci.* **2021**, *22*, 2060. [CrossRef] [PubMed]

33. Marshall, A.S.; Jones, N.S. Discovering Cellular Mitochondrial Heteroplasmy Heterogeneity with Single Cell RNA and ATAC Sequencing. *Biology* **2021**, *10*, 503. [CrossRef]

34. Macken, W.L.; Vandrovcova, J.; Hanna, M.G.; Pitceathly, R.D.S. Applying genomic and transcriptomic advances to mitochondrial medicine. *Nat. Rev. Neurol.* **2021**, *17*, 215–230. [CrossRef]

35. Poole, O.V.; Pizzamiglio, C.; Murphy, D.; Falabella, M.; Macken, W.L.; Bugiardini, E.; Woodward, C.E.; Labrum, R.; Efthymiou, S.; Salpietro, V.; et al. Mitochondrial DNA Analysis from Exome Sequencing Data Improves Diagnostic Yield in Neurological Diseases. *Ann. Neurol.* **2021**, *89*, 1240–1247. [CrossRef] [PubMed]

36. Lopes, L.R.; Murphy, D.; Bugiardini, E.; Salem, R.; Jager, J.; Futema, M.; Akhtar, M.M.; Savvatis, K.; Woodward, C.; Pittman, A.M.; et al. Iterative Reanalysis of Hypertrophic Cardiomyopathy Exome Data Reveals Causative Pathogenic Mitochondrial DNA Variants. *Circ-Genom. Precis. Me.* **2021**, *14*, 379–382. [CrossRef]

37. Gusic, M.; Prokisch, H. Genetic basis of mitochondrial diseases. *Febs. Lett.* **2021**, *595*, 1132–1158. [CrossRef] [PubMed]

38. Alston, C.L.; Stenton, S.L.; Hudson, G.; Prokisch, H.; Taylor, R.W. The genetics of mitochondrial disease: Dissecting mitochondrial pathology using multi-omic pipelines. *J. Pathol.* **2021**, *254*, 430–442. [CrossRef] [PubMed]

39. Rodriguez-Anaya, L.Z.; Felix-Sastre, A.J.; Lares-Villa, F.; Lares-Jimenez, L.F.; Gonzalez-Galaviz, J.R. Application of the omics sciences to the study of Naegleria fowleri, Acanthamoeba spp., and Balamuthia mandrillaris: Current status and future projections. *Parasite* **2021**, *28*, 36. [CrossRef]

40. Montarry, J.; Mimee, B.; Danchin, E.G.J.; Koutsovoulos, G.D.; Ste-Croix, D.T.; Grenier, E. Recent Advances in Population Genomics of Plant-Parasitic Nematodes. *Phytopathology* **2021**, *111*, 40–48. [CrossRef]

41. Stam, R.; Gladieux, P.; Vinatzer, B.A.; Goss, E.M.; Potnis, N.; Candresse, T.; Brewer, M.T. Population Genomic- and Phylogenomic-Enabled Advances to Increase Insight Into Pathogen Biology and Epidemiology Introduction. *Phytopathology* **2021**, *111*, 8–11. [CrossRef]

42. Ste-Croix, D.T.; St-Marseille, A.F.G.; Lord, E.; Belanger, R.R.; Brodeur, J.; Mimee, B. Genomic Profiling of Virulence in the Soybean Cyst Nematode Using Single-Nematode Sequencing. *Phytopathology* **2021**, *111*, 137–148. [CrossRef]

43. Brito, I.L. Examining horizontal gene transfer in microbial communities. *Nat. Rev. Microbiol.* **2021**, *19*, 442–453. [CrossRef] [PubMed]

44. Lyu, R.; Tsui, V.; McCarthy, D.J.; Crismani, W. Personalized genome structure via single gamete sequencing. *Genome Biol.* **2021**, *22*, 1–9. [CrossRef] [PubMed]

45. Campoy, J.A.; Sun, H.Q.; Goel, M.; Jiao, W.B.; Folz-Donahue, K.; Wang, N.; Rubio, M.; Liu, C.; Kukat, C.; Ruiz, D.; et al. Gamete binning: Chromosome-level and haplotype-resolved genome assembly enabled by high-throughput single-cell sequencing of gamete genomes. *Genome Biol.* **2020**, *21*, 1–20. [CrossRef]

46. Wang, C.R.; Liu, H.F.; Wang, H.Y.; Tao, J.J.; Yang, T.W.; Chen, H.; An, R.; Wang, J.; Huang, N.; Gong, X.Y.; et al. Robust Storage of Chinese Language in a Pool of Small Single-Stranded DNA Rings and Its Facile Reading-Out. *B Chem. Soc. Jpn.* **2021**, *94*, 53–59. [CrossRef]

47. Jha, A.B.; Gali, K.K.; Alam, Z.; Lachagari, V.B.R.; Warkentin, T.D. Potential Application of Genomic Technologies in Breeding for Fungal and Oomycete Disease Resistance in Pea. *Agronomy* **2021**, *11*, 1260. [CrossRef]

48. van Dijk, A.D.J.; Kootstra, G.; Kruijer, W.; de Ridder, D. Machine learning in plant science and plant breeding. *Iscience* **2021**, *24*, 101890. [CrossRef]

49. Awika, H.O.; Mishra, A.K.; Gill, H.; DiPiazza, J.; Avila, C.A.; Joshi, V. Selection of nitrogen responsive root architectural traits in spinach using machine learning and genetic correlations. *Sci. Rep.* **2021**, *11*, 1–13. [CrossRef]

50. Varshney, R.K.; Bohra, A.; Yu, J.M.; Graner, A.; Zhang, Q.F.; Sorrells, M.E. Feature Review Designing Future Crops: Genomics-Assisted Breeding Comes of Age. *Trends Plant Sci.* **2021**, *26*, 631–649. [CrossRef]

51. Anwar, K.; Joshi, R.; Dhankher, O.P.; Singla-Pareek, S.L.; Pareek, A. Elucidating the Response of Crop Plants towards Individual, Combined and Sequentially Occurring Abiotic Stresses. *Int. J. Mol. Sci.* **2021**, *22*, 6119. [CrossRef] [PubMed]

52. Pazhamala, L.T.; Kudapa, H.; Weckwerth, W.; Millar, A.H.; Varshney, R.K. Systems biology for crop improvement. *Plant Genome* **2021**, 1–23. [CrossRef]

53. Saad, N.S.M.; Severn-Ellis, A.A.; Pradhan, A.; Edwards, D.; Batley, J. Genomics Armed With Diversity Leads the Way in Brassica Improvement in a Changing Global Environment. *Front. Genet.* **2021**, *12*, 600789. [CrossRef]

54. Hu, D.D.; Jing, J.J.; Snowdon, R.J.; Mason, A.S.; Shen, J.X.; Meng, J.L.; Zou, J. Exploring the gene pool of Brassica napus by genomics-based approaches. *Plant Biotechnol. J.* **2021**. [CrossRef]

55. Witzel, K.; Kurina, A.B.; Artemyeva, A.M. Opening the Treasure Chest: The Current Status of Research on Brassica oleracea and B. rapa Vegetables From ex situ Germplasm Collections. *Front. Plant Sci.* **2021**, *12*, 925. [CrossRef] [PubMed]

56. Scossa, F.; Alseekh, S.; Fernie, A.R. Integrating multi-omics data for crop improvement. *J. Plant Physiol.* **2021**, *257*, 153352. [CrossRef]

57. Sinha, P.; Singh, V.K.; Bohra, A.; Kumar, A.; Reif, J.C.; Varshney, R.K. Genomics and breeding innovations for enhancing genetic gain for climate resilience and nutrition traits. *Theor. Appl. Genet.* **2021**, *134*, 1829–1843. [CrossRef]

58. Kuo, R.I.; Cheng, Y.Y.; Zhang, R.X.; Brown, J.W.S.; Smith, J.; Archibald, A.L.; Burt, D.W. Illuminating the dark side of the human transcriptome with long read transcript sequencing. *BMC Genom.* **2020**, *21*, 1–22. [CrossRef]

59. Sahlin, K.; Medvedev, P. Error correction enables use of Oxford Nanopore technology for reference-free transcriptome analysis (vol 12, 2, 2021). *Nat. Commun.* **2021**, *12*, 1–13. [CrossRef]

60. Liu, Z.Q.; Rabadan, R. Computing the Role of Alternative Splicing in Cancer. *Trends Cancer* **2021**, *7*, 347–358. [CrossRef]

61. Mertes, C.; Scheller, I.F.; Yepez, V.A.; Celik, M.H.; Liang, Y.J.Q.; Kremer, L.S.; Gusic, M.; Prokisch, H.; Gagneur, J. Detection of aberrant splicing events in RNA-seq data using FRASER. *Nat. Commun.* **2021**, *12*, 1–13. [CrossRef]

62. Hu, Y.; Fang, L.; Chen, X.L.; Zhong, J.F.; Li, M.Y.; Wang, K. LIQA: Long-read isoform quantification and analysis. *Genome Biol.* **2021**, *22*, 1–21. [CrossRef]

63. Riepe, T.V.; Khan, M.; Roosing, S.; Cremers, F.P.M.; 't Hoen, P.A.C. Benchmarking deep learning splice prediction tools using functional splice assays. *Hum. Mutat.* **2021**, *42*, 799–810. [CrossRef]

64. Pan, Y.; Kadash-Edmondson, K.E.; Wang, R.; Phillips, J.; Liu, S.; Ribas, A.; Aplenc, R.; Witte, O.N.; Xing, Y. RNA Dysregulation: An Expanding Source of Cancer Immunotherapy Targets. *Trends Pharmacol. Sci.* **2021**, *42*, 268–282. [CrossRef] [PubMed]

65. Macosko, E.Z.; Basu, A.; Satija, R.; Nemesh, J.; Shekhar, K.; Goldman, M.; Tirosh, I.; Bialas, A.R.; Kamitaki, N.; Martersteck, E.M.; et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **2015**, *161*, 1202–1214. [CrossRef]

66. Darmanis, S.; Sloan, S.A.; Zhang, Y.; Enge, M.; Caneda, C.; Shuer, L.M.; Gephart, M.G.H.; Barres, B.A.; Quake, S.R. A survey of human brain transcriptome diversity at the single cell level. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7285–7290. [CrossRef]

67. Kaster, A.K.; Sobol, M.S. Microbial single-cell omics: The crux of the matter. *Appl. Microbiol. Biot.* **2020**, *104*, 8209–8220. [CrossRef] [PubMed]

68. Adil, A.; Kumar, V.; Jan, A.T.; Asger, M. Single-Cell Transcriptomics: Current Methods and Challenges in Data Acquisition and Analysis. *Front. Neurosci-Switz* **2021**, *15*, 398. [CrossRef]

69. Song, Y.L.; Xu, X.; Wang, W.; Tian, T.; Zhu, Z.; Yang, C.Y. Single cell transcriptomics: Moving towards multi-omics. *Analyst* **2019**, *144*, 3172–3189. [CrossRef]

70. Grindberg, R.V.; Yee-Greenbaum, J.L.; McConnell, M.J.; Novotny, M.; O'Shaughnessy, A.L.; Lambert, G.M.; Arauzo-Bravo, M.J.; Lee, J.; Fishman, M.; Robbins, G.E.; et al. RNA-sequencing from single nuclei. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19802–19807. [CrossRef]

71. Hahn, O.; Fehlmann, T.; Zhang, H.; Munson, C.N.; Vest, R.T.; Borcherding, A.; Liu, S.; Villarosa, C.; Drmanac, S.; Drmanac, R.; et al. CooIMPS for robust sequencing of single-nuclear RNAs captured by droplet-based method. *Nucleic Acids Res.* **2021**, *49*, e11. [CrossRef]

72. Zhao, Z.H.; Ma, J.Y.; Meng, T.G.; Wang, Z.B.; Yue, W.; Zhou, Q.; Li, S.; Feng, X.; Hou, Y.; Schatten, H.; et al. Single-cell RNA sequencing reveals the landscape of early female germ cell development. *Faseb J.* **2020**, *34*, 12634–12645. [CrossRef]

73. Wen, L.; Tang, F.C. Human Germline Cell Development: From the Perspective of Single-Cell Sequencing. *Mol. Cell* **2019**, *76*, 320–328. [CrossRef]

74. Brandt, L.; Cristinelli, S.; Ciuffi, A. Single-Cell Analysis Reveals Heterogeneity of Virus Infection, Pathogenicity, and Host Responses: HIV as a Pioneering Example. *Annu. Rev. Virol.* **2020**, *7*, 333–350. [CrossRef] [PubMed]

75. Iqbal, F.; Lupieri, A.; Aikawa, M.; Aikawa, E. Harnessing Single-Cell RNA Sequencing to Better Understand How Diseased Cells Behave the Way They Do in Cardiovascular Disease. *Arterioscl. Throm. Vas. Biol.* **2021**, *41*, 585–600. [CrossRef]

76. Yu, S.G.; Li, C.H.; Lin, H.; Ou, M.L.; Tang, D.G.; Dai, Y.; Yan, Q. Application of single-cell RNA sequencing in embryonic development. *Genomics* **2020**, *112*, 4547–4551. [CrossRef]

77. Yasen, A.; Aini, A.; Wang, H.; Li, W.D.; Zhang, C.S.; Ran, B.; Tuxun, T.; Maimaitinijiati, Y.; Shao, Y.M.; Aji, T.; et al. Progress and applications of single-cell sequencing techniques. *Infect. Genet. Evol.* **2020**, *80*, 104198. [CrossRef] [PubMed]

78. Scatena, C.; Murtas, D.; Tomei, S. Cutaneous Melanoma Classification: The Importance of High-Throughput Genomic Technologies. *Front. Oncol.* **2021**, *11*, 635488. [CrossRef]

79. Zang, J.Y.; Ye, K.Y.; Fei, Y.; Zhang, R.Y.; Chen, H.G.; Zhuang, G.L. Immunotherapy in the Treatment of Urothelial Bladder Cancer: Insights From Single-Cell Analysis. *Front. Oncol.* **2021**, *11*, 2020. [CrossRef]

80. Wang, B.; Zhang, Y.Y.; Qing, T.; Xing, K.C.; Li, J.; Zhen, T.M.; Zhu, S.B.; Zhan, X.B. Comprehensive analysis of metastatic gastric cancer tumour cells using single-cell RNA-seq. *Sci. Rep.* **2021**, *11*, 1–10. [CrossRef]

81. Mercatelli, D.; Balboni, N.; Palma, A.; Aleo, E.; Sanna, P.P.; Perini, G.; Giorgi, F.M. Single-Cell Gene Network Analysis and Transcriptional Landscape of MYCN-Amplified Neuroblastoma Cell Lines. *Biomolecules* **2021**, *11*, 177. [CrossRef]

82. Ysebaert, L.; Quillet-Mary, A.; Tosolini, M.; Pont, F.; Laurent, C.; Fournie, J.J. Lymphoma Heterogeneity Unraveled by Single-Cell Transcriptomics. *Front. Immunol.* **2021**, *12*, 202. [CrossRef] [PubMed]

83. Lei, Y.L.; Tang, R.; Xu, J.; Wang, W.; Zhang, B.; Liu, J.; Yu, X.J.; Shi, S. Applications of single-cell sequencing in cancer research: Progress and perspectives. *J. Hematol. Oncol.* **2021**, *14*, 1–26. [CrossRef]

84. Liu, J.; Xu, T.M.; Jin, Y.M.; Huang, B.Y.; Zhang, Y. Progress and Clinical Application of Single-Cell Transcriptional Sequencing Technology in Cancer Research. *Front. Oncol.* **2021**, *10*, 3367. [CrossRef] [PubMed]

85. Wang, H.; Meng, D.; Guo, H.Y.; Sun, C.L.; Chen, P.X.; Jiang, M.L.; Xu, Y.; Yu, J.; Fang, Q.Y.; Zhu, J.; et al. Single-Cell Sequencing, an Advanced Technology in Lung Cancer Research. *Oncotargets Ther.* **2021**, *14*, 1895–1909. [CrossRef] [PubMed]

86. Guo, T.T.; Li, W.M.; Cai, X.Y. Applications of Single-Cell Omics to Dissect Tumor Microenvironment. *Front. Genet.* **2020**, *11*, 548719. [CrossRef]

87. Zhang, M.; Hu, S.F.; Min, M.; Ni, Y.L.; Lu, Z.; Sun, X.T.; Wu, J.Q.; Liu, B.; Ying, X.M.; Liu, Y. Dissecting transcriptional heterogeneity in primary gastric adenocarcinoma by single cell RNA sequencing. *Gut* **2021**, *70*, 464–475. [CrossRef]

88. Cheng, S.J.; Li, Z.Y.; Gao, R.R.; Xing, B.C.; Gao, Y.N.; Yang, Y.; Qin, S.S.; Zhang, L.; Ouyang, H.Q.; Du, P.; et al. A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell* **2021**, *184*, 792–809. [CrossRef]

89. Sinjab, A.; Han, G.C.; Wang, L.H.; Kadara, H. Field Carcinogenesis in Cancer Evolution: What the Cell Is Going On? *Cancer Res.* **2020**, *80*, 4888–4891. [CrossRef]

90. Cildir, G.; Yip, K.H.; Pant, H.; Tergaonkar, V.; Lopez, A.F.; Tumes, D.J. Understanding mast cell heterogeneity at single cell resolution. *Trends Immunol.* **2021**, *42*, 523–535. [CrossRef]

91. de Jong, E.; Bosco, A. Unlocking immune-mediated disease mechanisms with transcriptomics. *Biochem. Soc. Trans.* **2021**, *49*, 705–714. [CrossRef]

92. Derakhshani, A.; Rostami, Z.; Safarpour, H.; Shadbad, M.A.; Nourbakhsh, N.S.; Argentiero, A.; Taefehshokr, S.; Tabrizi, N.J.; Kooshkaki, O.; Astamal, R.V.; et al. From Oncogenic Signaling Pathways to Single-Cell Sequencing of Immune Cells: Changing the Landscape of Cancer Immunotherapy. *Molecules* **2021**, *26*, 2278. [CrossRef] [PubMed]

93. Guo, Y.E.; Li, Y.R.; Cai, B.L.; He, Q.Z.; Chen, G.F.; Wang, M.F.; Wang, K.; Wan, X.P.; Yan, Q. Phenotyping of immune and endometrial epithelial cells in endometrial carcinomas revealed by single-cell RNA sequencing. *Aging* **2021**, *13*, 6565–6591. [CrossRef] [PubMed]

94. Su, S.; Li, X.H. Dive into Single, Seek Out Multiple: Probing Cancer Metastases via Single-Cell Sequencing and Imaging Techniques. *Cancers* **2021**, *13*, 1067. [CrossRef] [PubMed]

95. Chen, Z.; Yu, M.C.; Yan, J.L.; Guo, L.; Zhang, B.; Liu, S.; Lei, J.; Zhang, W.T.; Zhou, B.H.; Gao, J.; et al. PNOC Expressed by B Cells in Cholangiocarcinoma Was Survival Related and LAIR2 Could Be a T Cell Exhaustion Biomarker in Tumor Microenvironment: Characterization of Immune Microenvironment Combining Single-Cell and Bulk Sequencing Technology. *Front. Immunol.* **2021**, *12*, 828. [CrossRef]

96. Ren, X.W.; Zhang, L.; Zhang, Y.Y.; Li, Z.Y.; Siemers, N.; Zhang, Z.M. Insights Gained from Single-Cell Analysis of Immune Cells in the Tumor Microenvironment. *Annu. Rev. Immunol.* **2021**, *39*, 583–609. [CrossRef] [PubMed]

97. Kashima, Y.; Togashi, Y.; Fukuoka, S.; Kamada, T.; Irie, T.; Suzuki, A.; Nakamura, Y.; Shitara, K.; Minamide, T.; Yoshida, T.; et al. Potentiality of multiple modalities for single-cell analyses to evaluate the tumor microenvironment in clinical specimens. *Sci. Rep.* **2021**, *11*, 1–11. [CrossRef] [PubMed]

98. Yuan, X.; Wang, J.X.; Huang, Y.X.; Shangguan, D.G.; Zhang, P. Single-Cell Profiling to Explore Immunological Heterogeneity of Tumor Microenvironment in Breast Cancer. *Front. Immunol.* **2021**, *12*, 471. [CrossRef]

99. Feng, B.H.; Hess, J. Immune-Related Mutational Landscape and Gene Signatures: Prognostic Value and Therapeutic Impact for Head and Neck Cancer. *Cancers* **2021**, *13*, 1162. [CrossRef] [PubMed]

100. Guruprasad, P.; Lee, Y.G.; Kim, K.H.; Ruella, M. The current landscape of single-cell transcriptomics for cancer immunotherapy. *J. Exp. Med.* **2021**, *218*, e20201574. [CrossRef]

101. Dai, Z.; Gu, X.Y.; Xiang, S.Y.; Gong, D.D.; Man, C.F.; Fan, Y. Research and application of single-cell sequencing in tumor heterogeneity and drug resistance of circulating tumor cells. *Biomark. Res.* **2020**, *8*, 1–8. [CrossRef]

102. Armand, E.J.; Li, J.H.; Xie, F.M.; Luo, C.Y.; Mukamel, E.A. Single-Cell Sequencing of Brain Cell Transcriptomes and Epigenomes. *Neuron* **2021**, *109*, 11–26. [CrossRef]

103. Iqbal, M.M.; Hurgobin, B.; Holme, A.L.; Appels, R.; Kaur, P. Status and Potential of Single-Cell Transcriptomics for Understanding Plant Development and Functional Biology. *Cytom. Part A* **2020**, *97*, 997–1006. [CrossRef]

104. Shaw, R.; Tian, X.; Xu, J. Single-Cell Transcriptome Analysis in Plants: Advances and Challenges. *Mol. Plant* **2021**, *14*, 115–126. [CrossRef]

105. Ma, S.X.; Lim, S.B. Single-Cell RNA Sequencing in Parkinson's Disease. *Biomedicines* **2021**, *9*, 368. [CrossRef] [PubMed]

106. Wang, J.X.; Ma, A.J.; Chang, Y.Z.; Gong, J.T.; Jiang, Y.X.; Qi, R.; Wang, C.K.; Fu, H.J.; Ma, Q.; Xu, D. scGNN is a novel graph neural network framework for single-cell RNA-Seq analyses. *Nat. Commun.* **2021**, *12*, 1–11. [CrossRef]

107. Raj, A.; van den Bogaard, P.; Rifkin, S.A.; van Oudenaarden, A.; Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods* **2008**, *5*, 877–879. [CrossRef] [PubMed]

108. Stahl, P.L.; Salmen, F.; Vickovic, S.; Lundmark, A.; Navarro, J.F.; Magnusson, J.; Giacomello, S.; Asp, M.; Westholm, J.O.; Huss, M.; et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **2016**, *353*, 78–82. [CrossRef] [PubMed]

109. Rodriques, S.G.; Stickels, R.R.; Goeva, A.; Martin, C.A.; Murray, E.; Vanderburg, C.R.; Welch, J.; Chen, L.L.M.; Chen, F.; Macosko, E.Z. Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **2019**, *363*, 1463–1467. [CrossRef]

110. Almeida, N.; Chung, M.W.H.; Drudi, E.M.; Engquist, E.N.; Hamrud, E.; Isaacson, A.; Tsang, V.S.K.; Watt, F.M.; Spagnoli, F.M. Employing core regulatory circuits to define cell identity. *EMBO J.* **2021**, *40*, e106785. [CrossRef] [PubMed]

111. Longo, S.K.; Guo, M.G.; Ji, A.L.; Khavari, P.A. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat. Rev. Genet.* **2021**, *6*, 1–18. [CrossRef]

112. Waylen, L.N.; Nim, H.T.; Martelotto, L.G.; Ramialison, M. From whole-mount to single-cell spatial assessment of gene expression in 3D. *Commun. Biol.* **2020**, *3*, 1–11. [CrossRef]

113. Savulescu, A.F.; Jacobs, C.; Negishi, Y.; Davignon, L.; Mhlanga, M.M. Pinpointing Cell Identity in Time and Space. *Front. Mol. Biosci.* **2020**, *7*, 209. [CrossRef] [PubMed]

114. Liao, J.; Lu, X.Y.; Shao, X.; Zhu, L.; Fan, X.H. Uncovering an Organ's Molecular Architecture at Single-Cell Resolution by Spatially Resolved Transcriptomics. *Trends Biotechnol.* **2021**, *39*, 43–58. [CrossRef]

115. Chen, Y.W.; Song, J.; Ruan, Q.Y.; Zeng, X.; Wu, L.L.; Cai, L.F.; Wang, X.Q.; Yang, C.Y. Single-Cell Sequencing Methodologies: From Transcriptome to Multi-Dimensional Measurement. *Small Methods* **2021**, *5*, 2100111. [CrossRef]

116. Troskie, R.L.; Jafrani, Y.; Mercer, T.R.; Ewing, A.D.; Faulkner, G.J.; Cheetham, S.W. Long-read cDNA sequencing identifies functional pseudogenes in the human transcriptome. *Genome Biol.* **2021**, *22*, 1–15. [CrossRef]

117. Jin, S.Q.; Guerrero-Juarez, C.F.; Zhang, L.H.; Chang, I.; Ramos, R.; Kuan, C.H.; Myung, P.; Plikus, M.V.; Nie, Q. Inference and analysis of cell-cell communication using CellChat. *Nat. Commun.* **2021**, *12*, 1–20. [CrossRef] [PubMed]

118. Turei, D.; Valdeolivas, A.; Gul, L.; Palacio-Escat, N.; Klein, M.; Ivanova, O.; Olbei, M.; Gabor, A.; Theis, F.; Korcsmaros, T.; et al. Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. *Mol. Syst. Biol.* **2021**, *17*, e9923. [CrossRef] [PubMed]

119. Ghoshdastider, U.; Rohatgi, N.; Naeini, M.M.; Baruah, P.; Revkov, E.; Guo, Y.A.; Rizzetto, S.; Wong, A.M.L.; Solai, S.; Nguyen, T.T.; et al. Pan-Cancer Analysis of Ligand-Receptor Cross-talk in the Tumor Microenvironment. *Cancer Res.* **2021**, *81*, 1802–1812. [CrossRef]

120. Bechtel, T.J.; Reyes-Robles, T.; Fadeyi, O.O.; Oslund, R.C. Strategies for monitoring cell-cell interactions. *Nat. Chem. Biol.* **2021**, *17*, 641–652. [CrossRef]

121. Armingol, E.; Officer, A.; Harismendy, O.; Lewis, N.E. Deciphering cell-cell interactions and communication from gene expression. *Nat. Rev. Genet.* **2021**, *22*, 71–88. [CrossRef]

122. Shao, X.; Lu, X.Y.; Liao, J.; Chen, H.J.; Fan, X.H. New avenues for systematically inferring cell-cell communication: Through single-cell transcriptomics data. *Protein Cell* **2020**, *11*, 866–880. [CrossRef] [PubMed]

123. Hoffmann, A.; Spengler, D. Single-Cell Transcriptomics Supports a Role of CHD8 in Autism. *Int. J. Mol. Sci.* **2021**, *22*, 3261. [CrossRef] [PubMed]

124. Xin, R.J.; Gao, Y.; Gao, Y.; Wang, R.; Kadash-Edmondson, K.E.; Liu, B.; Wang, Y.D.; Lin, L.; Xing, Y. isoCirc catalogs full-length circular RNA isoforms in human transcriptomes. *Nat. Commun.* **2021**, *12*, 1–11. [CrossRef] [PubMed]

125. Ilgisonis, E.; Vavilov, N.; Ponomarenko, E.; Lisitsa, A.; Poverennaya, E.; Zgoda, V.; Radko, S.; Archakov, A. Genome of the Single Human Chromosome 18 as a "Gold Standard" for Its Transcriptome. *Front. Genet.* **2021**, *12*, 958. [CrossRef]

126. Bobrovskikh, A.; Doroshkov, A.; Mazzoleni, S.; Carteni, F.; Giannino, F.; Zubairova, U. A Sight on Single-Cell Transcriptomics in Plants Through the Prism of Cell-Based Computational Modeling Approaches: Benefits and Challenges for Data Analysis. *Front. Genet.* **2021**, *12*, 771. [CrossRef]

127. Liu, J.J.; Fan, Z.W.; Zhao, W.L.; Zhou, X.B. Machine Intelligence in Single-Cell Data Analysis: Advances and New Challenges. *Front. Genet.* **2021**, *12*, 807. [CrossRef]

128. Xiang, R.Z.; Wang, W.C.; Yang, L.; Wang, S.Y.; Xu, C.H.; Chen, X.W. A Comparison for Dimensionality Reduction Methods of Single-Cell RNA-seq Data. *Front. Genet.* **2021**, *12*, 646936. [CrossRef] [PubMed]

129. Qin, Y.F.; Zhang, W.W.; Sun, X.Q.; Nan, S.W.; Wei, N.N.; Wu, H.J.; Zheng, X.Q. Deconvolution of heterogeneous tumor samples using partial reference signals. *Plos Comput. Biol.* **2020**, *16*, e1008452. [CrossRef]

130. Zhang, Y.R.; Ma, Y.L.; Huang, Y.K.; Zhang, Y.; Jiang, Q.; Zhou, M.; Su, J.Z. Benchmarking algorithms for pathway activity transformation of single-cell RNA-seq data. *Comput. Struct. Biotec.* **2020**, *18*, 2953–2961. [CrossRef]

131. Schlieben, L.D.; Prokisch, H.; Yepez, V.A. How Machine Learning and Statistical Models Advance Molecular Diagnostics of Rare Disorders Via Analysis of RNA Sequencing Data. *Front. Mol. Biosci.* **2021**, *8*, 647277. [CrossRef]

132. Seong, H.J.; Han, S.W.; Sul, W.J. Prokaryotic DNA methylation and its functional roles. *J. Microbiol.* **2021**, *59*, 242–248. [CrossRef]

133. Cao, B.; Wu, X.L.; Zhou, J.L.; Wu, H.; Liu, L.L.; Zhang, Q.H.; DeMott, M.S.; Gu, C.; Wang, L.R.; You, D.L.; et al. Nick-seq for single-nucleotide resolution genomic maps of DNA modifications and damage. *Nucleic Acids Res.* **2020**, *48*, 6715–6725. [CrossRef]

134. Wei, Y.; Huang, Q.Q.; Tian, X.H.; Zhang, M.M.; He, J.K.; Chen, X.X.; Chen, C.; Deng, Z.X.; Li, Z.Q.; Chen, S.; et al. Single-molecule optical mapping of the distribution of DNA phosphorothioate epigenetics. *Nucleic Acids Res.* **2021**, *49*, 3672–3680. [CrossRef]

135. Tourancheau, A.; Mead, E.A.; Zhang, X.S.; Fang, G. Discovering multiple types of DNA methylation from bacteria and microbiome using nanopore sequencing. *Nat. Methods* **2021**, *18*, 491–498. [CrossRef]

136. Beaulaurier, J.; Schadt, E.E.; Fang, G. Deciphering bacterial epigenomes using modern sequencing technologies. *Nat. Rev. Genet.* **2019**, *20*, 157–172. [CrossRef] [PubMed]

137. Mannweiler, O.; Pinto-Carbo, M.; Lardi, M.; Agnoli, K.; Eberl, L. Investigation of Burkholderia cepacia Complex Methylomes via Single-Molecule, Real-Time Sequencing and Mutant Analysis. *J. Bacteriol.* **2021**, *203*, e00683-20. [CrossRef]

138. Payelleville, A.; Brillard, J. Novel Identification of Bacterial Epigenetic Regulations Would Benefit From a Better Exploitation of Methylomic Data. *Front. Microbiol.* **2021**, *12*, 1205. [CrossRef] [PubMed]

139. Oliveira, P.H.; Fang, G. Conserved DNA Methyltransferases: A Window into Fundamental Mechanisms of Epigenetic Regulation in Bacteria. *Trends Microbiol.* **2021**, *29*, 28–40. [CrossRef] [PubMed]

140. Spadar, A.; Perdigao, J.; Phelan, J.; Charleston, J.; Modesto, A.; Elias, R.; de Sessions, P.F.; Hibberd, M.L.; Campino, S.; Duarte, A.; et al. Methylation analysis of Klebsiella pneumoniae from Portuguese hospitals. *Sci. Rep.* **2021**, *11*, 1–10. [CrossRef] [PubMed]

141. Morovic, W.; Budinoff, C.R. Epigenetics: A New Frontier in Probiotic Research. *Trends Microbiol.* **2021**, *29*, 117–126. [CrossRef]

142. Carter, M.Q.; Pham, A.; Huynh, S.; Parker, C.T.; Miller, A.; He, X.H.; Hu, B.; Chain, P.S.G. DNA adenine methylase, not the PstI restriction-modification system, regulates virulence gene expression in Shiga toxin-producing Escherichia coli. *Food Microbiol.* **2021**, *96*, 103722. [CrossRef] [PubMed]

143. Modlin, S.J.; Conkle-Gutierrez, D.; Kim, C.; Mitchell, S.N.; Morrissey, C.; Weinrick, B.C.; Jacobs, W.R.; Ramirez-Busby, S.M.; Hoffner, S.E.; Valafar, F. Drivers and sites of diversity in the DNA adenine methylomes of 93 Mycobacterium tuberculosis complex clinical isolates. *Elife* **2020**, *9*, e58542. [CrossRef]

144. Gaultney, R.A.; Vincent, A.T.; Lorioux, C.; Coppee, J.Y.; Sismeiro, O.; Varet, H.; Legendre, R.; Cockram, C.A.; Veyrier, F.J.; Picardeau, M. 4-Methylcytosine DNA modification is critical for global epigenetic regulation and virulence in the human pathogen Leptospira interrogans. *Nucleic Acids Res.* **2020**, *48*, 12102–12115. [CrossRef] [PubMed]

145. Allue-Guardia, A.; Garcia, J.I.; Torrelles, J.B. Evolution of Drug-Resistant Mycobacterium tuberculosis Strains and Their Adaptation to the Human Lung Environment. *Front. Microbiol.* **2021**, *12*, 137. [CrossRef] [PubMed]

146. Murphy, T.R.; Xiao, R.; Hamilton-Brehm, S.D. Hybrid genome de novo assembly with methylome analysis of the anaerobic thermophilic subsurface bacterium Thermanaerosceptrum fracticalcis strain DRI-13(T). *Bmc Genom.* **2021**, *22*, 1–16. [CrossRef]

147. Choi, W.L.; Mok, Y.G.; Huh, J.H. Application of 5-Methylcytosine DNA Glycosylase to the Quantitative Analysis of DNA Methylation. *Int. J. Mol. Sci.* **2021**, *22*, 1072. [CrossRef] [PubMed]

148. Usai, G.; Vangelisti, A.; Simoni, S.; Giordani, T.; Natali, L.; Cavallini, A.; Mascagni, F. DNA Modification Patterns within the Transposable Elements of the Fig (*Ficus carica* L.) Genome. *Plants* **2021**, *10*, 451. [CrossRef]

149. Liu, J.Z.; He, Z.H. Small DNA Methylation, Big Player in Plant Abiotic Stress Responses and Memory. *Front. Plant Sci.* **2020**, *11*, 1977. [CrossRef]

150. Rojas-Rojas, F.U.; Vega-Arreguin, J.C. Epigenetic insight into regulatory role of chromatin covalent modifications in lifecycle and virulence of Phytophthora. *Env. Microbiol. Rep.* **2021**, *13*, 445–457. [CrossRef]

151. Reva, O.N.; Larisa, S.A.; Mwakilili, A.D.; Tibuhwa, D.D.; Lyantagaye, S.; Chan, W.Y.; Lutz, S.; Ahrens, C.H.; Vater, J.; Borriss, R. Complete genome sequence and epigenetic profile of Bacillus velezensis UCMB5140 used for plant and crop protection in comparison with other plant-associated Bacillus strains. *Appl. Microbiol. Biot.* **2020**, *104*, 7643–7656. [CrossRef]

152. Ashe, A.; Colot, V.; Oldroyd, B.P. How does epigenetics influence the course of evolution? *Philos. T R Soc. B* **2021**, *376*, 20200111. [CrossRef]

153. Loughland, I.; Little, A.; Seebacher, F. DNA methyltransferase 3a mediates developmental thermal plasticity. *Bmc Biol.* **2021**, *19*, 1–11. [CrossRef] [PubMed]

154. Beck, D.; Ben Maamar, M.; Skinner, M.K. Genome-wide CpG density and DNA methylation analysis method (MeDIP, RRBS, and WGBS) comparisons. *Epigenetics* **2021**, *5*, 1–13. [CrossRef] [PubMed]

155. Paun, O.; Verhoeven, K.J.F.; Richards, C.L. Opportunities and limitations of reduced representation bisulfite sequencing in plant ecological epigenomics. *New Phytol.* **2019**, *221*, 738–742. [CrossRef]

156. Horemans, N.; Spurgeon, D.J.; Lecomte-Pradines, C.; Saenen, E.; Bradshaw, C.; Oughton, D.; Rasnaca, I.; Kamstra, J.H.; Adam-Guillermin, C. Current evidence for a role of epigenetic mechanisms in response to ionizing radiation in an ecotoxicological context. *Environ. Pollut.* **2019**, *251*, 469–483. [CrossRef]

157. Belli, M.; Tabocchini, M.A. Ionizing Radiation-Induced Epigenetic Modifications and Their Relevance to Radiation Protection. *Int. J. Mol. Sci.* **2020**, *21*, 5993. [CrossRef]

158. Cabrera-Licona, A.; Perez-Anorve, X.I.; Flores-Fortis, M.; del Moral-Hernandez, O.; Gonzalez-de la Rosa, C.H.; Suarez-Sanchez, R.; Chavez-Saldana, M.; Arechaga-Ocampo, E. Deciphering the epigenetic network in cancer radioresistance. *Radiother Oncol.* **2021**, *159*, 48–59. [CrossRef]

159. Schang, A.L.; Saberan-Djoneidi, D.; Mezger, V. The impact of epigenomic next-generation sequencing approaches on our understanding of neuropsychiatric disorders. *Clin. Genet.* **2018**, *93*, 467–480. [CrossRef] [PubMed]

160. Wagh, K.; Ishikawa, M.; Garcia, D.A.; Stavreva, D.A.; Upadhyaya, A.; Hager, G.L. Mechanical Regulation of Transcription: Recent Advances. *Trends Cell Biol.* **2021**, *31*, 457–472. [CrossRef]

161. Gorini, F.; Scala, G.; Cooke, M.S.; Majello, B.; Amente, S. Towards a comprehensive view of 8-oxo-7,8-dihydro-2′-deoxyguanosine: Highlighting the intertwined roles of DNA damage and epigenetics in genomic instability. *DNA Repair* **2021**, *97*, 103027. [CrossRef]

162. Scott, C.A.; Duryea, J.D.; MacKay, H.; Baker, M.S.; Laritsky, E.; Gunasekara, C.J.; Coarfa, C.; Waterland, R.A. Identification of cell type-specific methylation signals in bulk whole genome bisulfite sequencing data. *Genome Biol.* **2020**, *21*, 1–23. [CrossRef] [PubMed]

163. Men, S.; Yu, Y.Y. Prospects for Use of Single-Cell Sequencing to Assess DNA Methylation in Asthma. *Med. Sci. Monitor* **2020**, *26*, e925514-1. [CrossRef]

164. Kamies, R.; Martinez-Jimenez, C.P. Advances of single-cell genomics and epigenomics in human disease: Where are we now? *Mamm. Genome* **2020**, *31*, 170–180. [CrossRef]

165. Jonaitis, P.; Kupcinskas, L.; Kupcinskas, J. Molecular Alterations in Gastric Intestinal Metaplasia. *Int. J. Mol. Sci.* **2021**, *22*, 5758. [CrossRef] [PubMed]

166. Li, D.M.; Cheng, J.; Zhu, Z.A.; Catalfamo, M.; Goerlitz, D.; Lawless, O.J.; Tallon, L.; Sadzewicz, L.; Calderone, R.; Bellanti, J.A. Treg-inducing capacity of genomic DNA of Bifidobacterium longum subsp. infantis. *Allergy Asthma Proc.* **2020**, *41*, 372–385. [CrossRef]

167. Bellanti, J.A.; Li, D.M. T Regulatory Cells in Human Health and Diseases. In *Advances in Experimental Medicine and Biology*; Zheng, S.G., Ed.; Springer: Singapore, 2021; Volume 1278, pp. 1–302.

168. Trevino, A.E.; Sinnott-Armstrong, N.; Andersen, J.; Yoon, S.J.; Huber, N.; Pritchard, J.K.; Chang, H.Y.; Greenleaf, W.J.; Pasca, S.P. Chromatin accessibility dynamics in a model of human forebrain development. *Science* **2020**, *367*, eaay1645. [CrossRef] [PubMed]

169. Day, J.J.; Childs, D.; Guzman-Karlsson, M.C.; Kibe, M.; Moulden, J.; Song, E.; Tahir, A.; Sweatt, J.D. DNA methylation regulates associative reward learning. *Nat. Neurosci.* **2013**, *16*, 1445–1452. [CrossRef] [PubMed]

170. MacBean, L.F.; Smith, A.R.; Lunnon, K. Exploring Beyond the DNA Sequence: A Review of Epigenomic Studies of DNA and Histone Modifications in Dementia. *Curr. Genet. Med. Rep.* **2020**, *8*, 79–92. [CrossRef]

171. Perkovic, M.N.; Paska, A.V.; Konjevod, M.; Kouter, K.; Strac, D.S.; Erjavec, G.N.; Pivac, N. Epigenetics of Alzheimer's Disease. *Biomolecules* **2021**, *11*, 195. [CrossRef]

172. Zeng, S.Y.; Hua, Y.W.; Zhang, Y.; Liu, G.F.; Zhao, C.C. GLEANER: A web server for GermLine cycle Expression ANalysis and Epigenetic Roadmap visualization. *Bmc Bioinform.* **2021**, *22*, 1–13. [CrossRef] [PubMed]

173. Khaneghah, A.M.; Fakhri, Y.; Nematollahi, A.; Seilani, F.; Vasseghian, Y. The Concentration of Acrylamide in Different Food Products: A Global Systematic Review, Meta-Analysis, and Meta-Regression. *Food Rev. Int.* **2020**, *7*, 1–19. [CrossRef]

174. Seal, C.J.; de Mul, A.; Eisenbrand, G.; Haverkort, A.J.; Franke, K.; Lalljie, S.P.D.; Mykkaenen, H.; Reimerdes, E.; Scholz, G.; Somoza, V.; et al. Risk-benefit considerations of mitigation measures on acrylamide content of foods—A case study on potatoes, cereals and coffee. *Brit. J. Nutr.* **2008**, *99*, S1–S46. [CrossRef] [PubMed]

175. Non, A.L. Social epigenomics: Are we at an impasse? *Epigenomics* **2021**. [CrossRef]

176. Brasil, S.; Neves, C.J.; Rijoff, T.; Falcao, M.; Valadao, G.; Videira, P.A.; Ferreira, V.D. Artificial Intelligence in Epigenetic Studies: Shedding Light on Rare Diseases. *Front. Mol. Biosci.* **2021**, *8*, 314. [CrossRef] [PubMed]

177. Cho, J.C. Omics-based microbiome analysis in microbial ecology: From sequences to information. *J. Microbiol.* **2021**, *59*, 229–232. [CrossRef]

178. Davey, L.; Valdivia, R.H. Bacterial genetics and molecular pathogenesis in the age of high throughput DNA sequencing. *Curr. Opin. Microbiol.* **2020**, *54*, 59–66. [CrossRef] [PubMed]

179. Kristensen, J.M.; Singleton, C.; Clegg, L.A.; Petriglieri, F.; Nielsen, P.H. High Diversity and Functional Potential of Undescribed "Acidobacteriota" in Danish Wastewater Treatment Plants. *Front. Microbiol.* **2021**, *12*, 906. [CrossRef]

180. Singleton, C.M.; Petriglieri, F.; Kristensen, J.M.; Kirkegaard, R.H.; Michaelsen, T.Y.; Andersen, M.H.; Kondrotaite, Z.; Karst, S.M.; Dueholm, M.S.; Nielsen, P.H.; et al. Connecting structure to function with the recovery of over 1000 high-quality metagenome-assembled genomes from activated sludge using long-read sequencing. *Nat. Commun.* **2021**, *12*, 1–13. [CrossRef] [PubMed]

181. Infante-Villamil, S.; Huerlimann, R.; Jerry, D.R. Microbiome diversity and dysbiosis in aquaculture. *Rev. Aquacult.* **2021**, *13*, 1077–1096. [CrossRef]

182. Seeneevassen, L.; Bessede, E.; Megraud, F.; Lehours, P.; Dubus, P.; Varon, C. Gastric Cancer: Advances in Carcinogenesis Research and New Therapeutic Strategies. *Int. J. Mol. Sci.* **2021**, *22*, 3418. [CrossRef]

183. Nearing, J.T.; Comeau, A.M.; Langille, M.G.I. Identifying biases and their potential solutions in human microbiome studies. *Microbiome* **2021**, *9*, 1–22. [CrossRef]

184. Bokulich, N.A.; Ziemski, M.; Robeson, M.S.; Kaehler, B.D. Measuring the microbiome: Best practices for developing and benchmarking microbiomics methods. *Comput. Struct. Biotec.* **2020**, *18*, 4048–4062. [CrossRef]

185. Cusco, A.; Perez, D.; Vines, J.; Fabregas, N.; Francino, O. Long-read metagenomics retrieves complete single-contig bacterial genomes from canine feces. *Bmc Genom.* **2021**, *22*, 1–5. [CrossRef] [PubMed]

186. Zhang, M.; Dang, N.; Ren, D.Y.; Zhao, F.Y.; Lv, R.R.; Ma, T.; Bao, Q.H.; Menghe, B.; Liu, W.J. Comparison of Bacterial Microbiota in Raw 'Using PacBio Single Molecule Real-Time Sequencing Technology. *Front. Microbiol.* **2020**, *11*, 2708. [CrossRef]

187. Marco, M.L. Defining how microorganisms benefit human health. *Microb. Biotechnol.* **2021**, *14*, 35–40. [CrossRef] [PubMed]

188. Zhang, J.; Su, L.K.; Wang, Y.; Deng, S.L. Improved High-Throughput Sequencing of the Human Oral Microbiome: From Illumina to PacBio. *Can. J. Infect. Dis. Med.* **2020**, *2020*, 13. [CrossRef]

189. Xu, F.; Ge, C.T.; Li, S.T.; Tang, S.L.; Wu, X.W.; Luo, H.; Deng, X.Y.; Zhang, G.T.; Stevenson, A.; Baker, R.C. Evaluation of nanopore sequencing technology to identify Salmonella enterica Choleraesuis var. Kunzendorf and Orion var. 15(+), 34(+). *Int. J. Food Microbiol.* **2021**, *346*, 109167. [CrossRef] [PubMed]

190. Arumugam, K.; Bessarab, I.; Haryono, M.A.S.; Liu, X.H.; Zuniga-Montanez, R.E.; Roy, S.; Qiu, G.L.; Drautz-Moses, D.I.; Law, Y.Y.; Wuertz, S.; et al. Recovery of complete genomes and non-chromosomal replicons from activated sludge enrichment microbial communities with long read metagenome sequencing. *Npj Biofilms Microbi.* **2021**, *7*, 1–13. [CrossRef] [PubMed]

191. Lu, I.N.; Muller, C.P.; He, F.Q. Applying next-generation sequencing to unravel the mutational landscape in viral quasispecies. *Virus Res.* **2020**, *283*, 197963. [CrossRef] [PubMed]

192. He, W.Q.; Gao, Y.H.; Wen, Y.Q.; Ke, X.M.; Ou, Z.J.; Li, Y.Z.; He, H.; Chen, Q. Detection of Virus-Related Sequences Associated With Potential Etiologies of Hepatitis in Liver Tissue Samples From Rats, Mice, Shrews, and Bats. *Front. Microbiol.* **2021**, *12*, 1409. [CrossRef]

193. Ono, Y.; Asai, K.; Hamada, M. PBSIM2: A simulator for long-read sequencers with a novel generative model of quality scores. *Bioinformatics* **2021**, *37*, 589–595. [CrossRef] [PubMed]

194. Zablocki, O.; Michelsen, M.; Burris, M.; Solonenko, N.; Warwick-Dugdale, J.; Ghosh, R.; Pett-Ridge, J.; Sullivan, M.B.; Temperton, B. VirION2: A short- and long-read sequencing and informatics workflow to study the genomic diversity of viruses in nature. *Peerj* **2021**, *9*, e11088. [CrossRef]

195. Arbas, S.M.; Busi, S.B.; Queiros, P.; de Nies, L.; Herold, M.; May, P.; Wilmes, P.; Muller, E.E.L.; Narayanasamy, S. Challenges, Strategies, and Perspectives for Reference-Independent Longitudinal Multi-Omic Microbiome Studies. *Front. Genet.* **2021**, *12*, 858. [CrossRef]

196. Barry, G. Integrating the roles of long and small non-coding RNA in brain function and disease. *Mol. Psychiatr.* **2014**, *19*, 410–416. [CrossRef]

197. Guennewig, B.; Cooper, A.A. The Central Role of Noncoding RNA in the Brain. *Int. Rev. Neurobiol.* **2014**, *116*, 153–194. [CrossRef] [PubMed]

198. Muir, P.; Li, S.T.; Lou, S.K.; Wang, D.F.; Spakowicz, D.J.; Salichos, L.; Zhang, J.; Weinstock, G.M.; Isaacs, F.; Rozowsky, J.; et al. The real cost of sequencing: Scaling computation to keep pace with data generation. *Genome Biol.* **2016**, *17*, 1–9. [CrossRef]