HHS Public Access

Author manuscript

Neurosci Lett. Author manuscript; available in PMC 2017 December 29.

Published in final edited form as:

Neurosci Lett. 2016 June 20; 625: 47–55. doi:10.1016/j.neulet.2016.02.013.

Epigenetic regulation of mitochondrial function in neurodegenerative disease: new insights from advances in genomic technologies

Matthew Devall[1], **Janou Roubroeks**[1,2], **Jonathan Mill**[1,3], **Michael Weedon**[1], and **Katie Lunnon**[1,*]

[1]Institute of Clinical and Biomedical Science, University of Exeter Medical School, University of Exeter, Devon, UK [2]Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience (MHENS), Maastricht University, Maastricht, The Netherlands [3]Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College London, De Crespigny Park, London, UK

Abstract

The field of mitochondrial epigenetics has received increased attention in recent years and changes in mitochondrial DNA (mtDNA) methylation has been implicated in a number of diseases, including neurodegenerative diseases such as amyotrophic lateral sclerosis. However, current publications have been limited by the use of global or targeted methods of measuring DNA methylation. In this review, we discuss current findings in mitochondrial epigenetics as well as its potential role as a regulator of mitochondria within the brain. Finally, we summarize the current technologies best suited to capturing mtDNA methylation, and how a move towards whole epigenome sequencing of mtDNA may help to advance our current understanding of the field.

Keywords

Mitochondria; DNA methylation; Epigenetics; mtDNA; Alzheimer's disease; AD

Introduction

There is a resurgence of interest in the field of mitochondrial epigenetics, as mitochondrial dysfunction has been implicated in a variety of complex diseases including cancer [1, 2], amyotrophic lateral sclerosis [3, 4], and Alzheimer's disease (AD) [5]. Unlike modifications to nuclear DNA (ncDNA), which are now recognized as a mechanism through which the environment can influence biological processes and contribute to the development of a range

*Corresponding author: Dr Katie Lunnon, University of Exeter Medical School, RILD, Barrack Road, University of Exeter, Devon, UK. UK. Tel: + 44 1392 408 298. k.lunnon@exeter.ac.uk.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of different disease phenotypes [6-8], mtDNA methylation has only recently started to be investigated for its role in health and disease.

Despite this recent interest, few studies of mtDNA methylation have extended past global studies of mtDNA methylation using immunocytochemical techniques or the investigation of single candidate genes within the mitochondrial genome [9]. Throughout the course of this review we consider the caveats of studying epigenetic variation in mitochondria, which are particularly pertinent for studies of brain disorders given that changes in cell type composition and mtDNA copy number variation, which occur frequently in neurodegenerative diseases [10, 11], could have profound effects on the interpretation of findings in the field of mitochondrial epigenetics. Finally we consider how recent advances in genomic technologies could be applied to allow accurate and quantifiable measurements of the mitochondrial epigenome in the brain.

The Role of Mitochondria in the Brain

Mitochondria are the "powerhouse" of the cell and their dynamic nature allows for the organelle to be moved to areas of high ATP demand, for example cortical neurons require approximately 4.7 billion ATP per second to ensure continuous function [12]. Mitochondria are also implicated in many other important neurophysiological functions, for example synaptic mitochondria are believed to be involved in the regulation of neurotransmission by buffering extra intracellular Ca^{2+} [13], making the mitochondria a vital organelle in the establishment of Ca^{2+} homeostasis. Further, mitochondrial morphology and dynamics vary between synaptically immature and mature cortical neurons, with shorter mitochondrial lengths being observed in immature neurons allowing for increased movement and greater ability to meet the high energy demands of immature neurons, providing a key role for this organelle in neuronal development [14]. However, despite being the major site of reactive oxygen species (ROS) production, the mitochondria lack protective histones and defective mtDNA repair has been suggested to play a role in a number of neurodegenerative diseases [15]. As a by-product of ATP generation, ROS accumulation in areas of high ATP demand, such as post-mitotic neuronal cells, has been associated with neuronal loss [16]. Mutations in mitochondrial-encoded genes have been shown to cause a number of maternally heritable, monogenic diseases. Most notably, mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS) syndrome, a disease affecting multiple organs which can lead to the development of a number of syndromes, ranging from muscular weakness, fatigue and stroke-like episodes to dementia, epilepsy and diabetes in later stages [17]. As such, mtDNA variants have been shown to be an important driver of mitochondrial dysfunction and can have significant neuropathological effects [18].

MtDNA methylation and regulation of the mitochondria

The diverse range of mitochondrial functions coupled with their importance in the governance of cellular energy demands means that the expression of mitochondrial proteins requires sophisticated levels of fine-tuning. Although ~99% of mitochondrial proteins are encoded by the nuclear genome, the 16.569 Kb mitochondrial genome contains a total of 37 genes: 22 tRNAs, two rRNAs and 13 genes that encode for polypeptides along the electron transport chain (ETC) [19]. Within the nuclear genome, epigenetic processes mediate the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

reversible regulation of gene expression, occurring independently of DNA sequence, acting principally through chemical modifications to DNA and nucleosomal histone proteins. Epigenetic mechanisms orchestrate a diverse range of important neurobiological and cognitive processes in the brain and epigenetic modifications to the nuclear genome have been widely hypothesized to play a role in many neurological disorders, including Alzheimer's disease (AD) [11, 20, 21]. Given that mitochondrial dysfunction is a prominent feature of AD, we recently hypothesized that epigenetic modifications to the mitochondrial genome could be important in disease progression and pathology [5].

Despite rapid progress in the field of nuclear epigenetics, the field of mitochondrial epigenetics has received little attention since initial, but contradictory studies in the field were published in the 1970's [22-24]. These controversies continue, with one recent report concluding an absence of biologically significant levels of mtDNA methylation in four regions of mtDNA analyzed in human HEK293 cells and publically available data [25]. However, given the low sequencing depth of mtDNA in the publically available data used in this study (94×), and that multiple copies of the mtDNA genome are present in any given cell, it is possible that the true extent of mtDNA methylation was not determined. In contrary to this finding, both 5-methylcytosine (5-mC) and 5-hydroxymethylcystosine (5-hmC) have been identified in mammalian mitochondria from cell lines [26]. Further 5-mC has been found to co-localise in motor neurons with the mitochondrial marker superoxide dismutase 2 (SOD2).[4]. In addition, the identification of DNA methyltransferase 3a, (DNMT3a), an enzyme that plays an important role in catalyzing the transfer of the methyl moiety from the methyl donor protein S-adenoysl methionine (SAM), in mitochondria isolated from a mouse motor neuron cell line [4], and in mitochondria isolated from mouse brain and spinal cord [26], suggests potential tissue-specific mtDNA methylation. As the mitochondrial genome lacks protective histones, mtDNA instead being packaged via nucleoid proteins [27], post-translational histone modifications will not have a direct role on mitochondrial function. MtDNA methylation, on the other hand, may play an important role in regulating mitochondrial function, either independently, or via a complex mechanism involving nuclear-mitochondrial crosstalk [5]. Recent studies have suggested that mtDNA may have patterns of DNA methylation similar to that of plant and fungi, with methylation of cytosines adjacent to adenosine and thymine, in addition to the traditional CpG dinucleotide [28]. Since we recently suggested that mtDNA methylation could play a role in AD and other disorders characterized by mitochondrial dysfunction [5], further studies have been published to support our hypothesis. For instance, one study assessed mtDNA methylation in platelets in subjects with cardiovascular disease across seven genes using pyrosequencing and identified four genes with significantly increased levels of mtDNA methylation compared to controls [9]. In particular one gene, *MT-C01*, demonstrated >18% difference between groups. Another study showed differences in DNA methylation of the D-Loop and *MT-RNR1* were concordant with methylation differences in the subtelomeric region *D4Z4* in umbilical cord blood as well as the fetal and maternal side of the placenta, suggesting a potentially common epigenetic signature in these tissues [29]. The same group also recently demonstrated a positive correlation between airborne particulate matter exposure and placental mtDNA methylation in the D-Loop and *MT-RNR1* using targeted bisulfite-pyrosequencing [30]. This study also showed that increased mtDNA methylation is

Author Manuscript

accompanied by a reduction in mtDNA content, a potential marker of mitophagy, potentially indicating that alterations in mtDNA methylation of the D-Loop, a region highly involved in mtDNA transcription and replication, may lead to changes in mitochondrial biogenesis.

The role of mtDNA genetic variation on underlying mitochondrial function has become increasingly studied in recent years. MtDNA haplogroups and genetic variation have been found to have pronounced effects on mitochondrial functions, leading to altered ETC functions and ROS levels, which are associated with increased breast cancer risk [31]. However, given that a single mitochondrion can contain up to 10 copies of the mitochondrial genome [32], and as there are multiple mitochondria in any given cell, the extent of heteroplasmic mutations and their effect on mitochondrial function is far from clear [33]. Further, relatively few of the studies investigating mtDNA methylation have related identified changes to alterations in gene expression. One recent study showed significant hypermethylation of the D-Loop is negatively associated with expression of three mitochondrial-encoded genes (*CYTB, ND6, COXII*) in retinal microvessels derived from diabetic retinopathy donors [34]. However, the group failed to find significant differences in the methylation of *CYTB* and conclude that, given the importance of the D-Loop in replication and transcription of the mitochondrial genome, changes in D-Loop methylation may result in transcriptional changes across the mitochondrial genome and potentially contribute to the pathogenesis of diabetic retinopathy. However, a recent *in vivo* study showed D-loop methylation did not correlate with the expression of mitochondrial-encoded genes during inflammation [35]. Briefly, it was found that lipopolysaccharide (LPS)-mediated inflammation in mice led to significant changes in mtDNA transcript levels which were reversed upon treatment with α-Lipoic acid (LA), an agent with anti-inflammatory properties. However, despite these transcriptional changes, neither treatment altered D-Loop methylation. Taken together, these two studies suggest that mitochondrial transcription may be driven by mtDNA methylation of the D-Loop; however, this phenomenon is not the sole driver of mitochondrial gene expression. Further studies should potentially employ a more holistic approach to investigating the relationship of mtDNA methylation, gene expression and potential nuclear-encoded effectors of mtDNA transcription. Further, the use of whole-methylome sequencing of the mitochondria may aid in the elucidation of base-specific mtDNA methylation changes which may be important in gene expression.

The Bi-directionality of mitoepigenetics

Interestingly, a number of nuclear factors are also known to affect mitochondrial function. A recent study found DNA methylation levels in the nuclear-encoded, mitochondrial-specific transcription factor DNA polymerase gamma A (*POLGA*) inversely correlate with mtDNA copy number in pluripotent and multipotent cell types [36]. The removal of mtDNA [37] or changes in mtDNA haplogroups [38] have been found to be associated with differences in ncDNA methylation levels, suggesting an important interplay between the two genomes and a bi-directionality to mitoepigenetics [39].

Cytoplasmic hybrids (cybrids) are an important cell type for studying mitochondrial function that contain identical nuclei to the parent cell, but different mtDNA, allowing for a controlled investigation into the role of mtDNA variants. One study found that cybrids with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

mtDNA derived from haplogroup J had consistently higher levels of global ncDNA methylation than haplogroup H, suggesting that mtDNA variants may play an important role in influencing ncDNA methylation [40]. This difference in DNA methylation was accompanied by significant decreases in the expression of six of the 11 nuclear-encoded genes assayed that are involved in DNA methylation and acetylation processes, as well as significant increases in the transcription of *MAT2B* and *MBD4*, both of which are important for DNA methylation. Given the importance of these two processes in epigenomic regulation, this further illustrates how mtDNA genetic variation can impact on the nuclear genome. Another recent study investigating the mitochondrial 3243A>G heteroplasmy, which causes several clinical phenotypes including MELAS, showed that changes in heteroplasmy levels led to large and widespread changes in mitochondrial and transcriptional regulation, including changes in transcripts involved in DNA methylation and histone acetylation. Interestingly, high levels of this heteroplasmic variant (50-90%) resulted in ncDNA gene expression patterns that are similar to those seen in AD, Parkinson's disease and Huntington's disease. With alterations in ncDNA methylation in neurodegenerative disorders now well established [11, 41-44], future studies may aim to study mtDNA and ncDNA epigenomes in parallel to gain an understanding of the complete cellular epigenetic landscape.

Other mitochondrial DNA modifications

The majority of mtDNA methylation research has focussed on the DNA methylation mark 5-mC. However a number of other epigenetic marks such as 5-hmC, 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC) have also been identified. Despite this, functional roles of 5-caC and 5-fC have yet to be elucidated, however evidence suggests that they are involved in the Tet-mediated oxidation pathway, which converts 5-mC back to unmodified cytosine [45]. In a recent study, 5-fC and 5-caC, but not 5-hmC were found to lead to modest blocks in DNA transcription when mediated by either T7 RNA polymerase (T7 RNAP), or by human RNA polymerase II in an *in vitro* transcription assay, which was also replicated in 2 human cell lines [46]. Interestingly, T7 RNAP shares a high degree of homology with mitochondrial polymerases, and given the identification of 5-hmC and Tet proteins in mitochondria [26, 47], this study hypothesized a mechanism by which mtDNA demethylation products may be important for the regulation of mtDNA transcription. However, at present, no study has investigated 5-fC and 5-caC levels in mtDNA and further investigation into the modulation of mtDNA methylation needs to be undertaken.

Methodologies appropriate for epigenetic studies of mtDNA

Given the growing interest in studying epigenetic changes in mtDNA in a range of pathologies, there are a number of important considerations pertinent to this field of research. First, tissue-specific effects of DNA methylation mean that profiling a disease-relevant tissue is likely to be critical. Further, different cell types will have varying levels of mtDNA copy number and different DNA methylation levels dependent upon cellular requirements and localized environments [5]. As such, studies investigating mtDNA methylation should consider mtDNA copy number as a potential confounder to subsequent analysis. This is a pertinent issue in mtDNA methylation analysis given that a number of studies have implicated an association between mtDNA methylation status and mtDNA copy

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

number [48] and that tissue homogenates contain a number of different cell types, many with varying levels of mtDNA copy numbers. One potential way to solve this problem would be to first isolate different cell types from tissue homogenate using Fluorescence-activated cell sorting (FACS). However isolating sufficient levels of mitochondria from individual cell populations to yield sufficient mtDNA for methylation analyses seems unlikely with current technologies requiring high input amounts of DNA.

Another important consideration is the most appropriate method for studying mtDNA methylation. A host of different platforms have been developed that are capable of assessing DNA methylation in the nuclear genome. However, given the relative nascence of the field of mtDNA epigenetics, the appropriateness of the wide range of genomic technologies available has never been reviewed for its utility in this field. Advances in genomic technologies has seen ncDNA methylation studies move from profiling DNA methylation at a handful of candidate CpG sites to epigenome-wide association studies (EWAS). However, the current workhorse for such studies, the Illumina 450K methylation array, and its successor the Illumina EPIC methylation array, provide no coverage of the mitochondrial genome. Given that mtDNA methylation is now becoming increasingly investigated, an understanding of the technologies available to assess genome-wide changes in mitochondrial methylation is of upmost importance. An overview of genomic technologies that could be utilised to measure mtDNA methylation is provided in Table 1.

To capture an accurate and precise representation of potential mtDNA methylation changes, a technique should ideally be able to detect DNA modifications with low error rate and at single nucleotide resolution. This is more pertinent to studies of mtDNA epigenetics than in studies of ncDNA methylation given the low levels of mtDNA methylation identified in recent studies [30, 49]. Currently, the only published genome-wide study investigating mtDNA methylation utilized publically-available methylated DNA immunoprecipitation sequencing (MeDIP-Seq) data [50]. Whilst this study was able to identify spatio-temporal patterns of mtDNA methylation in a variety of cells and tissues, the coverage was relatively low (5×). Given sequencing and genomic biases, low levels of coverage may result in some bases not being sequenced adequately for analysis. Further, the resolution of MeDIP-Seq also remains a limitation. Despite high levels of concordance between MeDIP-Seq data, which averages the methylation values across a region, and other technologies that focus on single CpG sites [51], it is possible that a correlation may not be observed across every region of the mitochondrial genome. For example, averaging methylation in regions as small as 100bp in the mitochondrial genome could include averaging methylation across many gene boundaries of tRNAs in some regions and could greatly mask the true methylation status of these genes. Although this study was able to utilize publically available data that had been generated to inexpensively assess ncDNA methylation, studies to specifically assess mtDNA methylation would ideally investigate DNA methylation at single base resolution. Whole genome-bisulfite sequencing (WGB-Seq) is another source of publically available data that could be exploited to investigate mtDNA methylation, however although this allows the interrogation of DNA methylation at single nucleotide resolution, data with an appropriate depth of sequencing is lacking due to the expense of sequencing the entire ~3,000 Mb human genome. An alternative, and more appropriate next generation sequencing (NGS), approach would specifically capture the entire mitochondrial genome

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

prior to sequencing and due to its relatively small size would allow for inexpensive sequencing at single base resolution analysis. Currently, two custom capture methylation technologies are available from Agilent and Nimblegen, and coverage can be tailored to suit the researcher's needs, allowing for the interrogation of epigenetic variability within the mitochondrial genome. Despite costing considerably less than WGB-Seq, a targeted capture of the mitochondrial genome is still relatively expensive and requires a higher input of DNA.

Aside from cost versus coverage issues, sequencing the mitochondrial genome is associated with a number of unique caveats; most notably, the presence of nuclear mitochondria pseudogenes (*NUMT*s) in the nuclear genome. *NUMT*s are regions of the mitochondrial genome that, over an evolutionary period of time, have translocated to the nuclear genome and therefore share a high sequence homology with their mitochondrial paralogue. As such, failure to account for these regions has led to misinterpretations of genetic sequencing data [52, 53], and is one pitfall of using publically available data, although informatic approaches to account for these regions are being developed [50]. Ideally to avoid misinterpretation of data, mitochondria should be isolated from the tissue of choice prior to sequencing [5], and in epigenetic studies of brain disorders, this therefore requires the isolation of large quantities of mtDNA from small amounts of frozen post-mortem tissue [54].

Finally, the dependence on NGS technologies on PCR amplification prior to sequencing leads to a potential PCR bias and over-representation of some sequences. Given the tissue-specific nature of mitochondrial heteroplasmy, it is possible that these amplifications could influence findings. As such, a move towards third generation sequencing systems such as the Pac-Bio RS II and Oxford's Nanopore is an appealing prospect. Furthermore, despite both technologies being associated with a high individual base call error rate at present, the Pac-Bio's long read based deep sequencing has been used to accurately map complex regions of the human genome [55] as well as mitochondrial genomes in other species [56-58] [59]. Interestingly, given that both technologies use native DNA, both technologies provide the possibility of simultaneous genetic and epigenetic analysis, including a range of DNA modifications [60-64]. The circular consensus reads of PacBio RS II could provide an interesting platform for the simultaneous identification of the four major epigenetic marks in real time at a high coverage. However, isolation of high quality mtDNA from the range of frozen tissues commonly used in nuclear epigenetics could prove to be difficult and would require further investigation. The continual development of more accurate calling algorithms and likely future cost reductions may therefore lead to these third generation sequencing technologies being the optimal choice for providing specific, deep and targeted analysis of the mitochondrial (epi)genome.

Conclusions

Given the growing interest in studying mtDNA methylation in a range of brain disorders characterized by mitochondrial dysfunction, the advent of third generation sequencing technologies may allow for the accurate study of these small genomes in the near future. The circular consensus nature of Pac-Bio RS II potentially allows for the entire mitochondrial genome to be investigated from one long read, removing current NGS biases. However, this process would require isolation of intact mtDNA from a variety of potentially frozen and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

partially degraded tissues, which may prove challenging for researchers. Currently the use of NGS, particularly the more quantitative approaches of bisulfite-sequencing, may allow for the elucidation of mtDNA methylation and future studies should consider the utility of this approach versus candidate-based technologies such as pyrosequencing.

References

1. Feng S, et al. Correlation between increased ND2 expression and demethylated displacement loop of mtDNA in colorectal cancer. Mol Med Rep. 2012; 6(1):125–30. [PubMed: 22505229]
2. Wen SL, Zhang F, Feng S. Decreased copy number of mitochondrial DNA: A potential diagnostic criterion for gastric cancer. Oncol Lett. 2013; 6(4):1098–1102. [PubMed: 24137470]
3. Wong M, et al. Mitochondrial DNMT3A and DNA methylation in skeletal muscle and CNS of transgenic mouse models of ALS. Front Cell Neurosci. 2013; 7:279. [PubMed: 24399935]
4. Chestnut BA, et al. Epigenetic regulation of motor neuron cell death through DNA methylation. J Neurosci. 2011; 31(46):16619–36. [PubMed: 22090490]
5. Devall M, Mill J, Lunnon K. The mitochondrial epigenome: a role in Alzheimer's disease? Epigenomics. 2014; 6(6):665–75. [PubMed: 25531259]
6. Coppede F. The potential of epigenetic therapies in neurodegenerative diseases. Front Genet. 2014; 5:220. [PubMed: 25071843]
7. Rodriguez-Paredes M, Esteller M. Cancer epigenetics reaches mainstream oncology. Nat Med. 2011; 17(3):330–9. [PubMed: 21386836]
8. Robison AJ, Nestler EJ. Transcriptional and epigenetic mechanisms of addiction. Nat Rev Neurosci. 2011; 12(11):623–37. [PubMed: 21989194]
9. Baccarelli AA, Byun HM. Platelet mitochondrial DNA methylation: a potential new marker of cardiovascular disease. Clin Epigenetics. 2015; 7(1):44. [PubMed: 25901189]
10. Coskun P, et al. A mitochondrial etiology of Alzheimer and Parkinson disease. Biochim Biophys Acta. 2012; 1820(5):553–64. [PubMed: 21871538]
11. Lunnon K, Mill J. Epigenetic studies in Alzheimer's disease: Current findings, caveats, and considerations for future studies. Am J Med Genet B Neuropsychiatr Genet. 2013
12. Zhu XH, et al. Quantitative imaging of energy expenditure in human brain. Neuroimage. 2012; 60(4):2107–2117. [PubMed: 22487547]
13. Tang YG, Zucker RS. Mitochondrial involvement in post-tetanic potentiation of synaptic transmission. Neuron. 1997; 18(3):483–491. [PubMed: 9115741]
14. Chang DTW, Reynolds IJ. Differences in mitochondrial movement and morphology in young and mature primary cortical neurons in culture. Neuroscience. 2006; 141(2):727–736. [PubMed: 16797853]
15. Yang JL, et al. Mitochondrial DNA damage and repair in neurodegenerative disorders. DNA Repair. 2008; 7(7):1110–1120. [PubMed: 18463003]
16. Uttara B, et al. Oxidative Stress and Neurodegenerative Diseases: A Review of Upstream and Downstream Antioxidant Therapeutic Options. Current Neuropharmacology. 2009; 7(1):65–74. [PubMed: 19721819]
17. El-Hattab AW, et al. MELAS syndrome: Clinical manifestations, pathogenesis, and treatment options. Mol Genet Metab. 2015
18. Schaefer AM, et al. Prevalence of mitochondrial DNA disease in adults. Annals of Neurology. 2008; 63(1):35–39. [PubMed: 17886296]
19. Boengler K, Heusch G, Schulz R. Nuclear-encoded mitochondrial proteins and their role in cardioprotection. Biochimica Et Biophysica Acta-Molecular Cell Research. 2011; 1813(7):1286–1294.
20. Chouliaras L, et al. Epigenetic regulation in the pathophysiology of Alzheimer's disease. Prog Neurobiol. 2010; 90(4):498–510. [PubMed: 20097254]
21. Lardenoije R, et al. The epigenetics of aging and neurodegeneration. Prog Neurobiol. 2015

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

22. Dawid IB. 5-methylcytidylic acid: absence from mitochondrial DNA of frogs and HeLa cells. Science. 1974; 184(4132):80–1. [PubMed: 4815287]
23. Nass MM. Differential methylation of mitochondrial and nuclear DNA in cultured mouse, hamster and virus-transformed hamster cells. In vivo and in vitro methylation. J Mol Biol. 1973; 80(1): 155–75. [PubMed: 4361747]
24. Groot GS, Kroon AM. Mitochondrial DNA from various organisms does not contain internally methylated cytosine in -CCGG- sequences. Biochim Biophys Acta. 1979; 564(2):355–7. [PubMed: 158389]
25. Hong EE, et al. Regionally specific and genome-wide analyses conclusively demonstrate the absence of CpG methylation in human mitochondrial DNA. Mol Cell Biol. 2013; 33(14):2683–90. [PubMed: 23671186]
26. Shock LS, et al. DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. Proc Natl Acad Sci U S A. 2011; 108(9):3630–5. [PubMed: 21321201]
27. Kaufman BA, et al. The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. Mol Biol Cell. 2007; 18(9):3225–36. [PubMed: 17581862]
28. Bellizzi D, et al. The control region of mitochondrial DNA shows an unusual CpG and non-CpG methylation pattern. DNA Res. 2013; 20(6):537–47. [PubMed: 23804556]
29. Janssen BG, et al. Variation of DNA methylation in candidate age-related targets on the mitochondrial-telomere axis in cord blood and placenta. Placenta. 2014; 35(9):665–72. [PubMed: 25047690]
30. Janssen BG, et al. Placental mitochondrial methylation and exposure to airborne particulate matter in the early life environment: An ENVIRONAGE birth cohort study. Epigenetics. 2015; 10(6): 536–44. [PubMed: 25996590]
31. Bai RK, et al. Mitochondrial genetic background modifies breast cancer risk. Cancer Research. 2007; 67(10):4687–4694. [PubMed: 17510395]
32. Wallace DC. Mitochondrial-DNA Sequence Variation in Human-Evolution and Disease. Proceedings of the National Academy of Sciences of the United States of America. 1994; 91(19): 8739–8746. [PubMed: 8090716]
33. Wallace DC, Chalkia D. Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. Cold Spring Harb Perspect Med. 2013; 3(10):a021220.
34. Mishra M, Kowluru RA. Epigenetic Modification of Mitochondrial DNA in the Development of Diabetic Retinopathy. Invest Ophthalmol Vis Sci. 2015; 56(9):5133–42. [PubMed: 26241401]
35. Liu Z, et al. alpha-Lipoic acid attenuates LPS-induced liver injury by improving mitochondrial function in association with GR mitochondrial DNA occupancy. Biochimie. 2015
36. Lee W, et al. Mitochondrial DNA copy number is regulated by DNA methylation and demethylation of POLGA in stem and cancer cells and their differentiated progeny. Cell Death Dis. 2015; 6:e1664. [PubMed: 25719248]
37. Smiraglia DJ, et al. A novel role for mitochondria in regulating epigenetic modification in the nucleus. Cancer Biol Ther. 2008; 7(8):1182–90. [PubMed: 18458531]
38. Bellizzi D, et al. Global DNA methylation levels are modulated by mitochondrial DNA variants. Epigenomics. 2012; 4(1):17–27. [PubMed: 22332655]
39. Manev H, Dzitoyeva S, Chen H. Mitochondrial DNA: A Blind Spot in Neuroepigenetics. Biomol Concepts. 2012; 3(2):107–115. [PubMed: 22639700]
40. Atilano SR, et al. Mitochondrial DNA variants can mediate methylation status of inflammation, angiogenesis and signaling genes. Hum Mol Genet. 2015
41. Lunnon K, et al. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. Nat Neurosci. 2014; 17(9):1164–70. [PubMed: 25129077]
42. De Jager PL, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. Nat Neurosci. 2014; 17(9):1156–63. [PubMed: 25129075]
43. Ng CW, et al. Extensive changes in DNA methylation are associated with expression of mutant huntingtin. Proc Natl Acad Sci U S A. 2013; 110(6):2354–9. [PubMed: 23341638]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

44. Moore K, et al. Epigenome-wide association study for Parkinson's disease. Neuromolecular Med. 2014; 16(4):845–55. [PubMed: 25304910]
45. Song CX, He C. Potential functional roles of DNA demethylation intermediates. Trends Biochem Sci. 2013; 38(10):480–4. [PubMed: 23932479]
46. You C, et al. Effects of Tet-mediated oxidation products of 5-methylcytosine on DNA transcription in vitro and in mammalian cells. Sci Rep. 2014; 4:7052. [PubMed: 25394478]
47. Dzitoyeva S, Chen H, Manev H. Effect of aging on 5-hydroxymethylcytosine in brain mitochondria. Neurobiol Aging. 2012; 33(12):2881–91. [PubMed: 22445327]
48. Gao J, et al. De-methylation of displacement loop of mitochondrial DNA is associated with increased mitochondrial copy number and nicotinamide adenine dinucleotide subunit 2 expression in colorectal cancer. Mol Med Rep. 2015; 12(5):7033–8. [PubMed: 26323487]
49. Byun HM, et al. Effects of airborne pollutants on mitochondrial DNA methylation. Part Fibre Toxicol. 2013; 10:18. [PubMed: 23656717]
50. Ghosh S, Sengupta S, Scaria V. Comparative analysis of human mitochondrial methylomes shows distinct patterns of epigenetic regulation in mitochondria. Mitochondrion. 2014; 18:58–62. [PubMed: 25058022]
51. Clark C, et al. A comparison of the whole genome approach of MeDIP-seq to the targeted approach of the Infinium HumanMethylation450 BeadChip((R)) for methylome profiling. PLoS One. 2012; 7(11):e50233. [PubMed: 23209683]
52. Thangaraj K, et al. Sperm mitochondrial mutations as a cause of low sperm motility. J Androl. 2003; 24(3):388–92. [PubMed: 12721215]
53. Yao YG, et al. Pseudomitochondrial genome haunts disease studies. J Med Genet. 2008; 45(12): 769–72. [PubMed: 18611982]
54. Devall M, et al. A comparison of mitochondrial DNA isolation methods in frozen post-mortem human brain tissue—applications for studies of mitochondrial genetics in brain disorders. Biotechniques. 2015; 59(4):241–246. [PubMed: 26458552]
55. Larsen PA, Heilman AM, Yoder AD. The utility of PacBio circular consensus sequencing for characterizing complex gene families in non-model organisms. BMC Genomics. 2014; 15:720. [PubMed: 25159659]
56. Simpson L, et al. Comparison of the Mitochondrial Genomes and Steady State Transcriptomes of Two Strains of the Trypanosomatid Parasite, Leishmania tarentolae. PLoS Negl Trop Dis. 2015; 9(7):e0003841. [PubMed: 26204118]
57. Wolters JF, Chiu K, Fiumera HL. Population structure of mitochondrial genomes in Saccharomyces cerevisiae. BMC Genomics. 2015; 16:451. [PubMed: 26062918]
58. Taniguti LM, et al. Complete Genome Sequence of Sporisorium scitamineum and Biotrophic Interaction Transcriptome with Sugarcane. PLoS One. 2015; 10(6):e0129318. [PubMed: 26065709]
59. Peng Y, et al. De novo genome assembly of the economically important weed horseweed using integrated data from multiple sequencing platforms. Plant Physiol. 2014; 166(3):1241–54. [PubMed: 25209985]
60. Flusberg BA, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat Methods. 2010; 7(6):461–5. [PubMed: 20453866]
61. Powers JG, et al. Efficient and accurate whole genome assembly and methylome profiling of E. coli. BMC Genomics. 2013; 14:675. [PubMed: 24090403]
62. Lee WC, et al. The complete methylome of Helicobacter pylori UM032. BMC Genomics. 2015; 16:424. [PubMed: 26031894]
63. Schreiber J, et al. Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. Proc Natl Acad Sci U S A. 2013; 110(47): 18910–5. [PubMed: 24167260]
64. Laszlo AH, et al. Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. Proc Natl Acad Sci U S A. 2013; 110(47):18904–9. [PubMed: 24167255]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Highlights

- Variation in nuclear DNA methylation has been associated with neuropsychiatric and neurodegenerative diseases.
- The mitochondrial genome has been shown to be subject to regulation by DNA methylation.
- A potential role for alterations in mitochondrial DNA methylation in brain disorders has not yet been investigated.
- Next generation sequencing allows for an interrogation of the mitochondrial epigenome at single base resolution.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

An overwiew of sequencing technologies that could be utilized to study mtDNA methylation in neurodegenerative diseases

Generation	Method	General Overview	Advantages	Disadvantages
Next	**MeDIP- & hMeDIP-Seq**	Methylated DNA Immunoprecipitation sequencing (MeDIP-Seq) is an immunoprecipitation based method which uses monoclonal antibodies against 5-mC. In brief, purified genomic DNA is sheared by sonication to produce random fragments. These fragments are then denatured and immunoprecipitated, followed by PCR amplification. Using high-throughput sequencing at a depth of two Gigabases, around 70-80% of CpGs in the human genome can be identified [1].	**1** Staining with anti-5-hmC antibodies allows for the analysis 5-hmC Stroud et al. [2], which is not only present in the brain and significantly reduced in AD [3], but is also present in brain mitochondria [4] **2** Sequencing produces files typically less than 6GB (fasta format), which make alignments and bioinformatic analysis less computationally exhaustive. **3** Cost-effective relative to most bisulfite based approaches [5]. **4** Use of antibody means that only regions of interest are investigated.	**1** Investigation of both 5-mC and 5-hmC would be costly and would require several micrograms of mtDNA as typical hMeDIP- Seq experiments require 4 – 5 μg of DNA [2,6]. **2** Requirement of antibody binding reduces resolution to methylated windows. Analysis of single cytosine methylation sites is not possible [7]. **3** Count based data means that reads must be normalised for CpG density and total read counts [8], although pipelines are now more capable of accounting for this [9].
	Custom Capture	Custom capture kits available allow for a custom design of a library for the enrichment of specific DNA target regions. In brief, DNA undergoes standard NGS pipelines approaches of sonication, end repair, A-tailing and in this case, methyl-adaptor ligation before being bound by custom baits, bisulfite treated and amplified. Samples can then be run on a sequencer such as the Illumina HiSeq.	**1** Single base resolution for DNA methylation analysis **2** Customized library to capture 100% of mitochondrial genome leads to further mtDNA enrichment and reducing risk of *NUMT* amplification.	**1** Currently the process is not supported by Agilent. **2** Typically cost exceeds that of non-bisulfite experiments. **3** Reduced complexity of sequence due to bisulfite conversion [10].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Generation	Method	General Overview	Advantages	Disadvantages
			3 Can use as little as 1μg mtDNA (Nimblegen). **4** Small target genome would likely generate small output files from sequencing and could allow for rapid, less computationally exhaustive alignment. **5** Simultaneous analysis of CpG and non-CpG methylation. **6** Small size of the genome allows for ample space on a Mi-Seq lane for high (>1000×) coverage of over 100 samples at a time, potentially important given the multi-copy nature of the genome.	
	RRBS & ERRBS	Reduced Representation BS-Seq (RRBS) makes use of MspI restriction enzymes for selective digestion of genomic DNA. This produces fragments of genomic regions enriched for CpG sites, which can be bisulfite treated and sequenced irrespective of their methylation status. This way, many samples can be processed efficiently and inexpensively [11].	**1** Requires much less DNA input than WGS-BS, and most other techniques, with RRBS requiring as little as 10ng [12] and Enhanced RRBS requiring around 50ng or less [13] **2** Can provide coverage of CpG promoters and CpG islands at a fraction of the cost of WGS-BS. **3** Single base resolution **4** Has been modified to allow for single cell epigenomic analysis [14,15].	**1** Targets CpG promoters and CpG islands, however the mitochondrial genome contains no CpG islands, therefore could give a very poor coverage of the mitochondrial methylome. **2** Reduced complexity of sequence due to bisulfite conversion [10].
	WGS-BS & OxBS-Seq	The whole genome shotgun bisulfite sequencing (WGS-BS) technique converts	**1** Single base resolution for DNA	**1** Sequences whole genome, including

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Generation	Method	General Overview	Advantages	Disadvantages
		fragmented DNA cytosines to uracil bases by sodium bisulfite treatment, while (hydroxy)methylcytosines remain unmodified. After PCR amplification, methylcytosines will be read as thymine in a sequencer, and can thus be distinguished from unmethylated sites. The reads can then be aligned to recreate the DNA sequence [16]. The recent development of the Oxidative Bisulfite (OxBS) method allows for quantification of hydroxymethylation by converting a 5hmC base to 5fC before bisulfite treatment. This enables a direct measurement of 5mC, and an indirect measurement of 5hmC at base-pair resolution [17].	methylation analysis **2** Simultaneous analysis of CpG and non-CpG methylation **3** The development of OxBS-Seq allows for the identification of both 5-hmC and 5-mC when used in combination with standard BS-Seq [18,19]. **4** OxBS-seq can be further combined with reduced Bisulfite Sequencing to allow for a characterisation of 5-mC, 5-hmC and 5-formylcytosine [20], allowing for an in depth characterization of the mitochondrial methylome.	many repetitive, non-methylated AT-rich regions. **2** Vast quantities of data generated leads to massive (100GB+ files) following sequencing, leading to time consuming processing [21]. **3** High cost of sequencing to improve sensitivity makes it unsuitable for large studies [10]. **4** Reduced complexity of sequence due to bisulfite conversion [10]
Generation 2.5	**PacBio**	Single-molecule real-time sequencing (SMRT) by Pacific Biosciences is achieved using zero-mode waveguide (ZMW) array technology. This technique uses a single DNA polymerase molecule attached to the bottom of a ZMW hole (tens of nanometers in diameter). By illuminating only the bottom 30nm of the ZMW with a laser, single nucleotide addition to the DNA can be measured. Each nucleotide fluoresces when bound to the DNA polymerase, which is detected by a camera before being cleaved off. Bases can be identified by corresponding fluorescent colours [22].	**1** Single molecule resolution technology avoids potential PCR bias by using pre-amplified DNA as an input for sequencing [23]. **2** Optimised for circular genomes, making the potential use of mtDNA analysis an interesting possibility. **3** Increased library complexity to bisulfite based methods as DNA modifications are determined based upon changes to polymerase kinetics specific to modification present [24]. Potentially, this	**1** The RSII platform is capable of producing continuous long reads and circular consensus reads, at present, these reads are typically associated with high error rates [28]. However, given the random nature of the errors, with sufficient depth SMRT can give an accurate read of the genome if sequenced at high coverage [29]. **2** The cost and size of PacBio RSII, coupled with the need for specialist technical and

Generation	Method	General Overview	Advantages	Disadvantages
			may be developed to accurately characterise all methylation marks in one simultaneous run. **4** Longer read length could lead to improved coverage and accuracy for processes such as variant calling [25]. **5** Ideal for de novo sequencing as capable of reads at a length of 8kbp [26] **6** Can detect strand specific patterns of 5hmC, after selective chemical labelling, in a high throughput manner [27]	bioinformatic support may limit much research to outsourcing data to specialist sequencing services.
Third Generation	**Nanopore**	The nanopore technique makes use of either biological nanopores or solid state nanopores, which are embedded in a membrane immersed in salt solution. To create a flow of ions through the pore, an electrical current can be applied. A single stranded DNA molecule passing through the pore will create measurable changes in the intensity of the current, and different DNA bases could be distinguished by the degree and duration of modulation of the current [30].	**1** Although still in alpha testing, nanopore technologies such as Oxford's MinION™ allow for sequencing analysis to be done on a portable, benchtop device in real-time [31]. **2** Identification of 5-hmC and 5mC **3** Systems like the Oxford MinION™ allows for the activation of hundreds of nanopores in parallel to reduce time of sequencing large genomes versus single pore technologies. **4** Sample preparation	**1** Observed per base error rates for applications such as DNA sequencing have reduced in recent years, but are still very high, recently reported at 30% [31]. **2** Some nanopore technologies utilise the binding of MBD1 to 5mC, altering current flow through the nanopore, to detect the methylation mark [33]; MBD1 has not yet been identified in mitochondria. **3** Although nanopore sequencing could in principle

Generation	Method	General Overview	Advantages	Disadvantages
			steps are reduced as no amplification or cloning is necessary, and use of enzymes is limited. This in turn reduces costs of sequencing [32].	require less than 1 µg of genomic DNA, an estimated 700 µg is needed due to the concentration-limited rate at which DNA molecules can be captured [32]

Abbreviations: Methylated DNA Immunoprecipitation sequencing (MeDIP-Seq), Hydroxymethylated DNA Immunoprecipitation sequencing (hMeDIP-Seq), Nuclear mitochondria pseudogenes (NUMTs), Next Generation Sequencing (NGS), Reduced Representation Bisulfite Sequencing (RRBS), Enhanced Reduced Representation Bisulfite Sequencing (ERRBS), Whole Genome Shotgun Bisulfite Sequencing (WGS-BS), 5-hydroxymethylcytosine (5-hmC), -methylcytosine (5-mC), Single-molecule real-time sequencing (SMRT), Zero-mode Waveguide (ZMW), Methyl DNA binding protein 1 (MBD1)

[1] Hirst, M. and Marra, M.A. (2010) Next generation sequencing based approaches to epigenomics. *Briefings in functional genomics*, **9**, 455-465.

[2] Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S. and Jacobsen, S.E. (2011) 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome biology*, **12**, R54.

[3] Condliffe, D., Wong, A., Troakes, C., Proitsi, P., Patel, Y., Chouliaras, L., Fernandes, C., Cooper, J., Lovestone, S., Schalkwyk, L. *et al.* (2014) Cross-region reduction in 5-hydroxymethylcytosine in Alzheimer's disease brain. *Neurobiology of aging*, **35**, 1850-1854.

[4] Dzitoyeva, S., Chen, H. and Manev, H. (2012) Effect of aging on 5-hydroxymethylcytosine in brain mitochondria. *Neurobiology of aging*, **33**, 2881-2891.

[5] Fouse, S.D., Nagarajan, R.O. and Costello, J.F. (2010) Genome-scale DNA methylation analysis. *Epigenomics*, **2**, 105-117.

[6] Tan, L., Xiong, L., Xu, W., Wu, F., Huang, N., Xu, Y., Kong, L., Zheng, L., Schwartz, L., Shi, Y. *et al.* (2013) Genome-wide comparison of DNA hydroxymethylation in mouse embryonic stem cells and neural progenitor cells by a new comparative hMeDIP-seq method. *Nucleic acids research*, **41**, e84.

[7] Huang, J., Renault, V., Sengenes, J., Touleimat, N., Michel, S., Lathrop, M. and Tost, J. (2012) MeQA: a pipeline for MeDIP-seq data quality assessment and analysis. *Bioinformatics*, **28**, 587-588.

[8] Laird, P.W. (2010) Principles and challenges of genomewide DNA methylation analysis. *Nature reviews. Genetics*, **11**, 191-203.

[9] Chavez, L., Jozefczuk, J., Grimm, C., Dietrich, J., Timmermann, B., Lehrach, H., Herwig, R. and Adjaye, J. (2010) Computational analysis of genome-wide DNA methylation during the differentiation of human embryonic stem cells along the endodermal lineage. *Genome research*, **20**, 1441-1450.

[10] Lee, E.J., Luo, J., Wilson, J.M. and Shi, H. (2013) Analyzing the cancer methylome through targeted bisulfite sequencing. *Cancer letters*, **340**, 171-178.

[11] Gu, H., Smith, Z.D., Bock, C., Boyle, P., Gnirke, A. and Meissner, A. (2011) Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nature protocols*, **6**, 468-481.

[12] Stirzaker, C., Taberlay, P.C., Statham, A.L. and Clark, S.J. (2014) Mining cancer methylomes: prospects and challenges. *Trends in genetics : TIG*, **30**, 75-84.

[13] Garrett-Bakelman, F.E., Sheridan, C.K., Kacmarczyk, T.J., Ishii, J., Betel, D., Alonso, A., Mason, C.E., Figueroa, M.E. and Melnick, A.M. (2015) Enhanced reduced representation bisulfite sequencing for assessment of DNA methylation at base pair resolution. *Journal of visualized experiments : JoVE*.

[14] Guo, H., Zhu, P., Wu, X., Li, X., Wen, L. and Tang, F. (2013) Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome research*, **23**, 2126-2135.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[15] Guo, H., Zhu, P., Guo, F., Li, X., Wu, X., Fan, X., Wen, L. and Tang, F. (2015) Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nature protocols*, **10**, 645-659.

[16] Krueger, F., Kreck, B., Franke, A. and Andrews, S.R. (2012) DNA methylome analysis using short bisulfite sequencing data. *Nat Meth*, **9**, 145-151.

[17] Stewart, S.K., Morris, T.J., Guilhamon, P., Bulstrode, H., Bachman, M., Balasubramanian, S. and Beck, S. (2015) oxBS-450K: A method for analysing hydroxymethylation using 450K BeadChips. *Methods*, **72**, 9-15.

[18] Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W. and Balasubramanian, S. (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science*, **336**, 934-937.

[19] Booth, M.J., Ost, T.W., Beraldi, D., Bell, N.M., Branco, M.R., Reik, W. and Balasubramanian, S. (2013) Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. *Nature protocols*, **8**, 1841-1851.

[20] Booth, M.J., Marsico, G., Bachman, M., Beraldi, D. and Balasubramanian, S. (2014) Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nature chemistry*, **6**, 435-440.

[21] Xi, Y. and Li, W. (2009) BSMAP: whole genome bisulfite sequence MAPping program. *BMC bioinformatics*, **10**, 232.

[22] Schadt, E.E., Turner, S. and Kasarskis, A. (2010) A window into third-generation sequencing. *Human molecular genetics*, **19**, R227-R240.

[23] Davis, B.M., Chao, M.C. and Waldor, M.K. (2013) Entering the era of bacterial epigenomics with single molecule real time DNA sequencing. *Current opinion in microbiology*, **16**, 192-198.

[24] Flusberg, B.A., Webster, D.R., Lee, J.H., Travers, K.J., Olivares, E.C., Clark, T.A., Korlach, J. and Turner, S.W. (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature methods*, **7**, 461-465.

[25] Sims, D., Sudbery, I., Ilott, N.E., Heger, A. and Ponting, C.P. (2014) Sequencing depth and coverage: key considerations in genomic analyses. *Nature reviews. Genetics*, **15**, 121-132.

[26] Steinbock, L. and Radenovic, A. (2015) The emergence of nanopores in next-generation sequencing. *Nanotechnology*, **26**, 074003.

[27] Song, C.X., Clark, T.A., Lu, X.Y., Kislyuk, A., Dai, Q., Turner, S.W., He, C. and Korlach, J. (2012) Sensitive and specific single-molecule sequencing of 5-hydroxymethylcytosine. *Nature methods*, **9**, 75-77.

[28] Ono, Y., Asai, K. and Hamada, M. (2013) PBSIM: PacBio reads simulator--toward accurate genome assembly. *Bioinformatics*, **29**, 119-121.

[29] Roberts, R.J., Carneiro, M.O. and Schatz, M.C. (2013) The advantages of SMRT sequencing. *Genome Biol*, **14**, 405.

[30] McGinn, S. and Gut, I.G. (2013) DNA sequencing–spanning the generations. *New biotechnology*, **30**, 366-372.

[31] Kilianski, A., Haas, J.L., Corriveau, E.J., Liem, A.T., Willis, K.L., Kadavy, D.R., Rosenzweig, C.N. and Minot, S.S. (2015) Bacterial and viral identification and differentiation by amplicon sequencing on the MinION nanopore sequencer. *Gigascience*, **4**, 12.

[32] Branton, D., Deamer, D.W., Marziali, A., Bayley, H., Benner, S.A., Butler, T., Di Ventra, M., Garaj, S., Hibbs, A. and Huang, X. (2008) The potential and challenges of nanopore sequencing. *Nature biotechnology*, **26**, 1146-1153.

[33] Hyun, B.R., McElwee, J.L. and Soloway, P.D. (2015) Single molecule and single cell epigenomics. *Methods*, **72**, 41-50.