Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

*Curr Opin Neurol*. Author manuscript; available in PMC 2017 October 01.

Published in final edited form as:

*Curr Opin Neurol*. 2016 October ; 29(5): 527–536. doi:10.1097/WCO.0000000000000374.

# Next generation sequencing in neuromuscular diseases

**S Efthymiou**[1,2], **A Manole**[1], and **H Houlden**[1]

[1]Department of Molecular Neuroscience and MRC Centre for Neuromuscular diseases, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK

[2]Department of Clinical and Experimental Epilepsy, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK

## Abstract

**Purpose of review—**Neuromuscular diseases are clinically and genetically heterogeneous and probably contains the greatest proportion of causative Mendelian defects than any other group of conditions. These disorders affect muscle and/or nerves with neonatal, childhood or adulthood onset, with significant disability and early mortality. Along with heterogeneity, unidentified and often very large genes, require complementary and comprehensive methods in routine molecular diagnosis. Inevitably this leads to increased diagnostic delays and challenges in the interpretation of genetic variants.

**Recent findings—**The application of next-generation sequencing, as a research and diagnostic strategy has made significant progress into solving many of these problems. The analysis of these data is by no means simple and the clinical input is essential to interpret results.

**Summary—**In this review, we describe using examples the recent advances in the genetic diagnosis of neuromuscular disorders, in research and clinical practice and the latest developments that are underway in NGS. We also discuss the latest collaborative initiatives such as the Genomics England genome sequencing project that combine rare disease clinical phenotyping with genomics, with the aim of defining the vast majority of rare disease genes in patients as well as modifying risks and pharmacogenomics factors.

### Keywords

exome; genome; NGS; nerve; neuromuscular

## Introduction

Neuromuscular disorders (NMD) are a broad group of conditions that affect muscles (myopathies, dystrophies, ion channel diseases and malignant hyperthermia) (1–6), nerves (Charcot Marie Tooth (CMT) (7, 8), motor neurone disease (MND) (9–13) hereditary spastic paraplegia (14–18) and spinal muscular atrophies (19–21) and neuromuscular junctions (myasthenic syndromes) (22). Muscle weakness, wasting, fasciculation, cramps, numbness,

Contact: s.efthymiou@ucl.ac.uk; telephone: +44 (0) 203448 4069.

**Conflicts of interest**

There are no conflicts of interest.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

respiratory and cranial nerve palsies are common features in several NMDs. These disorders are frequently inherited and extremely heterogeneous with more than 500 implicated genes. There are many subgroups but individually they are rare and often severe, affecting a wide age group from children to adults (23–28).

The advance in gene discovery can be attributed to progress in technology development and affordability of next-generation sequencing (NGS); moving from PCR in the 1980s to exome (protein-coding sequences) and gene panels in 2009/10 to the emerging use of genome sequencing in 2015 (29) (Figure 1A and Table 1 and 2). With size and coverage come difficulties with bioinformatics and variant interpretation (Figure 1B). Given the extreme heterogeneity of NMDs and the large number of disease genes, they are well suited for the use of NGS (23). Here we also discuss the challenges of NGS and how the introduction of NGS into research and diagnostic clinical practice has changed the management of NMD patients (Table 1).

## Research and Diagnostic Challenges of NGS

NMD genetic heterogeneity is present in the group as a whole, as well as sub-groups. For example in autosomal recessive limb girdle muscular dystrophy there are over 20 genes implicated and few diagnostic clues to narrow the genetic causes down. The aim of genetic testing is to make a molecular diagnosis but genetic analysis alone is not sufficient to prove a particular genetic variant, the skill of the referring clinician is essential in assessing variants with the clinical differential diagnosis. A diagnosis of an inherited NMD may seem clear in the context of a positive family history or when multiple siblings are affected but the mode of inheritance is often difficult to define in small families or where family history may be lacking, as in apparently ‘sporadic’ cases where the mode of inheritance can be any pattern or *de-novo* dominant. The lack of family history and indistinct NMD phenotypes makes filtering of NGS data difficult, often requiring other unaffected family members to help with genetic proof and the need for further functional laboratory work as proof (30, 31).

In the pre-NGS era, we and others have published screening and guidelines for genetic testing of NMD such as CMT (7) that were based on careful phenotyping and sequential Sanger sequencing of the most prevalent candidate genes. Although achieving a genetic diagnosis in over 60% of patients with CMT, Sanger sequencing of the ever increasing number of individual genes is no longer cost or time effective. In CMT type 1, over 70% of patients have the 17p duplication and over 80% of patients receive a diagnosis using traditional methods but in CMT2, distal HMN and HSN, where 40–80% of disease-causing genes are unknown the diagnostic rates are only between 17 and 30% (8). With the advancement in gene discovery and NGS, we would expect these figures to be over 90% when all known and novel genetic causes could be analysed (32).

## Large NMD genes and different types of genetic defect

NMD genes are often large and contain many polymorphisms making the identification of disease causing variants difficult. This is particularly true for myopathies where genes such as Duchene muscular dystrophy (DMD) (MIM#300377) spanning more than 2.3 Mb with 79

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

exons, the titin gene (TTN) (MIM#188840) with 363 exons and an open reading frame spanning over 100 Kb and nebulin gene (NEB) (MIM #161650) with 183 exons may be defective. Different types of defects are detected in NMD genes (23). For example, 60% of patients with Duchenne and Becker muscular dystrophies have deletions in DMD, 10% have duplications, and the rest have point mutations or small insertions–deletions. Prior to NGS, Sanger sequencing, multiple ligation probe analysis (MLPA) and cytogenetic genomic hybridization (CGH) arrays would often be used to target mutation or deletion hot spots and therefore using unbiased NGS we will likely increase the frequency of detection of defects in these large genes. Although genome sequencing is capable of identifying deletions and duplications, the frequency of these defects in Duchenne and Becker muscular dystrophies may mean that MLPA/CGH array is an easier and more cost effective approach.

## Next Generation sequencing

There are a number of challenges in moving from simple Sanger sequencing to the NGS analysis of multiple and eventually all the genes in the genome, aimed at revealing the vast majority of genetic variants. The advent of NGS has in the last decade led to an era of inexpensive, high-throughput DNA sequencing of large numbers of genes in a single reaction, thus facilitating the discovery of novel disease genes and variants (6, 10, 33–35).

### Whole exome sequencing (WES)

Whole exome sequencing (WES) is an efficient technology that can increase the diagnostic yield when searching for alleles causing rare Mendelian disorders (36) (37). Exome analyses the protein's encoding region where an estimated 85% of disease-causing mutations are believed to occur. Exomes are enriched by targeted hybridization and then each enriched genomic library is sequenced on a platform such as the Illumina Genome Analyzer II. This procedure covers around 96% of the targeted, mappable bases comprising the exomes of affected individuals. Filtering using databases such as dbSNP (http://www.ncbi.nlm.nih.gov/SNP/) and 1000 genomes http://www.1000genomes.org/ removes common variants, aiding in identifying the causal gene (Figures 2 and 3). It has been an invaluable tool in gene discovery, with around 100 genes linked with neuromuscular disorders discovered in this way (38) such as the identification of ADCK3 mutations using exome sequencing as shown in Figure 3 and 5. However, less-reachable regions of the genome with low sequence complexity restrict the ability to design useful WES capture baits, and differences in the hybridization efficiency of WES capture probes can result in off target capture effects as well as regions of the genome such as in exon 1, or GC-rich regions, with little or no coverage (33, 37, 39–41).

### Gene panels

Gene panels use next-generation sequencing to target genes and regions thought to be relevant to various diseases. Both custom-made panels that can include several hundred genes of interest and off-the-shelf panels of over 6000 genes are available. A custom-made panel with around 160 genes is used by the Diagnostics lab at the Institute of Neurology, and the same panel can be reduced to 40 genes when investigating neuropathies as well as when detecting mitochondrial deletions (Figure 4). Off-the-shelf panels such as Agilent focused

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

exome and the Illumina Trusight panels offer larger-scale screening of genes that are thought important such as all those on the Online Mendelian Inheritance in Man (OMIM) database. They rely on solution hybridization, in which pools of biotinylated oligonucleotide probes, each targeting a segment of the desired capture region, are used to pull down particular pieces of the genome. Gene panels are a quick relatively low-cost screening method to look for a few gene variants associated with specific NMDs.

NGS technologies have also advanced the genetic diagnosis of mitochondrial disorders in sequencing the mitochondrial genome or nuclear genes (42). A two-step strategy based on a targeted custom-made mitochondrial gene panel with 132 genes, and WES, was used by Legati *et al.* in a cohort of 125 early onset mitochondrial cases. However, an uncontroversial genetic diagnosis was only possible in 19 patients by targeted gene panel sequencing and in 6 out of 10 by WES. The rest of the cases although initially classified as mitochondria-disease patients, might be carriers of non-mitochondrial gene mutations (34).

Targeted NGS can be beneficial for diagnosing patients with early presentation of NMDs. Chae *et al.* used targeted NGS to screen 43 patients presented with early onset neuromuscular disorders of an unknown genetic origin. They were tested by panel for 579 nuclear genes associated with myopathy. Sanger sequencing has been the basic tool used for detecting mutations in myopathies but this is a laborious and expensive technique to be used for large genes. Compared to exome sequencing which is often more costly, has a higher false positive rate and turnaround time, and can be more difficult to interpret, targeted NGS is an attractive solution. Chae *et al.* were able to diagnose 21 of the 43 patients with a definite genetic cause for myopathy (23). In a similar way, Evila *et al.* used a custom-made panel, MyoCap, with 180 myopathy related genes, and succeeded in diagnosing 21 out of 61 patients, with 9 of them having potential disease-causing mutations in TTN (3).

### Whole genome sequencing (WGS)

Whole genome sequencing (WGS) allows examination of single-nucleotide variants (SNVs), indels, structural variants (SVs) and copy number variants (CNVs) in both the ~1% part of the genome that encodes protein sequences and the ~99% of remaining non-coding sequences. Therefore, WGS has more reliable sequence coverage with more uniformity. With the start of the 100,000 Project by the NHS in 2012, UCL has become a Genomics Medical Centre where patients with rare, life threatening and debilitating diseases such as many NMDs will be sequenced. Genomics England Clinical Interpretation Partnership (GeCIP) has also been established and aims to enhance clinical interpretation and validation of whole genome sequencing data delivered by the 100,000 genomes project (figure 6). It is likely to reveal many new variants and genes and derive new scientific and clinical findings. With the rapid drop in sequencing cost and the ability of WGS to rapidly produce large volumes of data, it is becoming a powerful tool for genomic research. However, WGS has still not been introduced in daily routine diagnostics, as it still remains more expensive and is more computationally demanding. Both WES and WGS generate vast amounts of complex data depicting the genetic variants, which require intensive filtering and complex interpretation (Table 2).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

## NGS limitations and improvements for clinical use

As NGS is employed both in research and diagnostic laboratories, some issues need to be addressed. One challenge of NGS is the detection of different types of mutations, especially repeat expansions or structural variations. Some of the most common NMDs are due to repeat expansions; these mutations may be missed is larger than the read length of 150bp, leading to false negative results. Increasing sequence coverage, longer reads, improving bioinformatics algorithms, software and novel sequencing technologies may solve this problem in order to propose an exclusion diagnosis (43). Alternatively, a combination of methods can be used. Detection of structural variations will benefit from whole genome sequencing (44–47) (table 1). Another obstacle is the incomplete coverage of commercial exome capture libraries. Although new versions of capture kits are released consistently, none capture all coding parts of all genes. This issue can be addressed by targeted sequencing of regions of interest and improving the capturing process.

### False positives

Another drawback is the high error rates in NGS compared to Sanger sequencing. Artificial mutations can be produced during template amplification or sequencing, leading to false positive results. Thus, Sanger sequencing of interesting variants detected by NGS is essential in validation but increases cost and turn-around time. This issue can be addressed by improving capture and sequencing approaches to increase variant coverage, achieving more reliable data. In addition, better data filtering protocols can reduce the pool of false positives (43).

### Volume of data

NGS generates a high volume of data which becomes problematic for analysis and storage.

Cloud computing can be a solution for reducing the cost of expensive computing infrastructure but needs to be secure. Importantly, decreasing costs in NGS has outpaced the increase in calculation power and storage capacity of computers. As sequencing becomes cheaper than data storage of the corresponding sequence output, re-sequencing of a patient DNA might be more cost-effective than saving original data (48).

### Variants filtering and mutation identification

Large numbers of variants are detected in NGS, it is difficult to distinguish between individual, rare and non-pathogenic variations without clinical significance versus disease-causing mutations. Defining the complete list of polymorphisms in different populations will require sequencing of world population groups. A recent study hypothesised that 27 % of published mutations appear to be sequencing errors, common polymorphisms, or have a lack of evidence of pathogenicity (45). This issue can be addressed by further analysis and validation such as *in-silico* prediction, detection of mutations in the same gene in unrelated individuals with the same disorder, absence in a control population, co-segregation in affected families and functional studies. Multiple criteria should be combined in order to prove the pathogenicity of the variation. In addition incidental findings can be problematic, especially for unreported diseases. Guidelines should be established and implemented by

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

national committees, as in the case of the Genomics England sequencing project in order to determine whether the analysis and reports should be selective or complete (30).

### Benefits of definitive molecular diagnosis

Identification of disease-causing mutations is vital for patients and families to provide an accurate diagnosis. The knowledge of the mutation and mutated gene improves disease management, and allows for inclusion into therapeutic trials. Genetic counseling is then possible, as carrier status determination and prenatal diagnosis can decrease the risk of recurrence (49). In addition, finding the disease-causing mutation permits potential phenotype–genotype correlations and greater understanding of the pathophysiological mechanisms, disease pathways and development of potential therapies (50, 51).

## Conclusion

There are limitations and challenges in using NGS but this technology is significantly developed for research and routine diagnostic practice. As the cost of NGS declines and the technology improves leading to obtaining faster and more accurate data, it is strongly anticipated that this recent technology will complement many clinical and pathological investigations to enable diagnosis as well as the identification of disease modifying and pharmacogenomics factors.

## Acknowledgments

We would like to thank Dr Maria Athanasopoulou for her assistance with the figures.

**Financial support and sponsorship**

This study was supported by the Medical Research Council, The Wellcome Trust, The Brain Research Trust (BRT) and the NINDS at the NIH. This study was also supported by the NIHR University College London Hospitals (UCLH) and Biomedical Research Centre (BRC).

## References and recommended reading

**Papers of particular interest, published within the annual period of review have been highlighted as:**

*Special interest

**Of outstanding interest

1*. Bucelli RC, Arhzaouy K, Pestronk A, Pittman SK, Rojas L, Sue CM, et al. SQSTM1 splice site mutation in distal myopathy with rimmed vacuoles. Neurology. 2015; 85(8):665–74. Original clinical description. [PubMed: 26208961]

2. Cortese A, Machado P, Morrow J, Dewar L, Hiscock A, Miller A, et al. Longitudinal observational study of sporadic inclusion body myositis: implications for clinical trials. Neuromuscular disorders: NMD. 2013; 23(5):404–12. [PubMed: 23489664]

3*. Evila A, Arumilli M, Udd B, Hackman P. Targeted next-generation sequencing assay for detection of mutations in primary myopathies. Neuromuscular disorders: NMD. 2016; 26(1):7–15. Exome sequencing in a series of myopathies. [PubMed: 26627873]

4. Hermans MC, Faber CG, Bekkers SC, de Die-Smulders CE, Gerrits MM, Merkies IS, et al. Structural and functional cardiac changes in myotonic dystrophy type 1: a cardiovascular magnetic

resonance study. Journal of Cardiovascular Magnetic Resonance. 2012; 14(1):48. [PubMed: 22827830]

5. Hunter JM, Ahearn ME, Balak CD, Liang WS, Kurdoglu A, Corneveaux JJ, et al. Novel pathogenic variants and genes for myopathies identified by whole exome sequencing. Molecular genetics & genomic medicine. 2015; 3(4):283–301. [PubMed: 26247046]

6*. Izumi R, Warita H, Niihori T, Takahashi T, Tateyama M, Suzuki N, et al. Isolated inclusion body myopathy caused by a multisystem proteinopathy-linked hnRNPA1 mutation. Neurology Genetics. 2015; 1(3):e23. Original clinical description. [PubMed: 27066560]

7. Murphy SM, Laura M, Fawcett K, Pandraud A, Liu YT, Davidson GL, et al. Charcot-Marie-Tooth disease: frequency of genetic subtypes and guidelines for genetic testing. Journal of neurology, neurosurgery, and psychiatry. 2012; 83(7):706–10.

8**. Rossor AM, Evans MR, Reilly MM. A practical approach to the genetic neuropathies. Practical neurology. 2015; 15(3):187–98. Guidelines for investigating neuropathies. [PubMed: 25898997]

9. Blokhuis AM, Groen EJ, Koppers M, van den Berg LH, Pasterkamp RJ. Protein aggregation in amyotrophic lateral sclerosis. Acta neuropathologica. 2013; 125(6):777–94. [PubMed: 23673820]

10. Borghero G, Pugliatti M, Marrosu F, Marrosu MG, Murru MR, Floris G, et al. TBK1 is associated with ALS and ALS-FTD in Sardinian patients. Neurobiology of aging. 2016

11. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron. 2011; 72(2):245–56. [PubMed: 21944778]

12. Gitler AD, Shorter J. RNA-binding proteins with prion-like domains in ALS and FTLD-U. Prion. 2011; 5(3):179–87. [PubMed: 21847013]

13**. Lim L, Wei Y, Lu Y, Song J. ALS-Causing Mutations Significantly Perturb the Self-Assembly and Interaction with Nucleic Acid of the Intrinsically Disordered Prion-Like Domain of TDP-43. PLoS biology. 2016; 14(1):e1002338. Functional work on an original clinical description of ALS. [PubMed: 26735904]

14*. Castro-Fernandez C, Arias M, Blanco-Arias P, Santome-Collazo L, Amigo J, Carracedo A, et al. Targeted NGS meets expert clinical characterization: Efficient diagnosis of spastic paraplegia type 11. Applied & translational genomics. 2015; 5:33–6. Original clinical description of the *SPG11* gene. [PubMed: 26937357]

15*. Lynch DS, Koutsis G, Tucci A, Panas M, Baklou M, Breza M, et al. Hereditary spastic paraplegia in Greece: characterisation of a previously unexplored population using next-generation sequencing. European journal of human genetics: EJHG. 2015 Original clinical description of the *SPG11* gene.

16. Rinaldi C, Schmidt T, Situ AJ, Johnson JO, Lee PR, Chen KL, et al. Mutation in CPT1C Associated With Pure Autosomal Dominant Spastic Paraplegia. JAMA neurology. 2015; 72(5): 561–70. [PubMed: 25751282]

17. Salinas S, Proukakis C, Crosby A, Warner TT. Hereditary spastic paraplegia: clinical features and pathogenetic mechanisms. The Lancet Neurology. 2008; 7(12):1127–38. [PubMed: 19007737]

18. Stevanin G, Santorelli FM, Azzedine H, Coutinho P, Chomilier J, Denora PS, et al. Mutations in SPG11, encoding spatacsin, are a major cause of spastic paraplegia with thin corpus callosum. Nature genetics. 2007; 39(3):366–72. [PubMed: 17322883]

19. Filla A, De Michele G. Overview of autosomal recessive ataxias. Handbook of clinical neurology. 2012; 103:265–74. [PubMed: 21827894]

20. Maksemous N, Roy B, Smith RA, Griffiths LR. Next-generation sequencing identifies novel CACNA1A gene mutations in episodic ataxia type 2. Molecular genetics & genomic medicine. 2016; 4(2):211–22. [PubMed: 27066515]

21. Synofzik M, Smets K, Mallaret M, Di Bella D, Gallenmuller C, Baets J, et al. SYNE1 ataxia is a common recessive ataxia with major non-cerebellar features: a large scale multi-centre study. Brain: a journal of neurology. 2016

22. Neuromuscular Disease Center. Available from: http://neuromuscular.wustl.edu/

23**. Chae JH, Vasta V, Cho A, Lim BC, Zhang Q, Eun SH, et al. Utility of next generation sequencing in genetic diagnosis of early onset neuromuscular disorders. Journal of medical



Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

genetics. 2015; 52(3):208–16. Exome sequencing in a series of neuromuscular disorders. [PubMed: 25635128]

24. Lindquist SG, Moller LB, Dali CI, Marner L, Kamsteeg EJ, Nielsen JE, et al. A Novel TTBK2 De Novo Mutation in a Danish Family with Early-Onset Spinocerebellar Ataxia. Cerebellum (London, England). 2016

25. Siekierska A, Isrie M, Liu Y, Scheldeman C, Vanthillo N, Lagae L, et al. Gain-of-function FHF1 mutation causes early-onset epileptic encephalopathy with cerebellar atrophy. Neurology. 2016

26. Tzoulis C, Sztromwasser P, Johansson S, Gjerde IO, Knappskog P, Bindoff LA. PNKP Mutations Identified by Whole-Exome Sequencing in a Norwegian Patient with Sporadic Ataxia and Edema. Cerebellum (London, England). 2016

27*. van de Warrenburg BP, Schouten MI, de Bot ST, Vermeer S, Meijer R, Pennings M, et al. Clinical exome sequencing for cerebellar ataxia and spastic paraplegia uncovers novel gene-disease associations and unanticipated rare disorders. European journal of human genetics: EJHG. 2016 Exome sequencing in a series of *SPG11* cases.

28. Wang W, Wang C, Dawson DB, Thorland EC, Lundquist PA, Eckloff BW, et al. Target-enrichment sequencing and copy number evaluation in inherited polyneuropathy. Neurology. 2016; 86(19): 1762–71. [PubMed: 27164712]

29. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. 1977. Biotechnology (Reading, Mass). 1992; 24:104–8.

30. MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, et al. Guidelines for investigating causality of sequence variants in human disease. Nature. 2014; 508(7497):469–76. [PubMed: 24759409]

31. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature. 2009; 461(7265):747–53. [PubMed: 19812666]

32. von Bubnoff A. Next-generation sequencing: the race is on. Cell. 2008; 132(5):721–3. [PubMed: 18329356]

33. Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. The New England journal of medicine. 2013; 369(16):1502–11. [PubMed: 24088041]

34**. Legati A, Reyes A, Nasca A, Invernizzi F, Lamantea E, Tiranti V, et al. New genes and pathomechanisms in mitochondrial disorders unraveled by NGS technologies. Biochimica et biophysica acta. 2016 Large series of mitochondrial cases investigated by NGS.

35. Endo Y, Dong M, Noguchi S, Ogawa M, Hayashi YK, Kuru S, et al. Milder forms of muscular dystrophy associated with POMGNT2 mutations. Neurology Genetics. 2015; 1(4):e33. [PubMed: 27066570]

36. Alazami AM, Patel N, Shamseldin HE, Anazi S, Al-Dosari MS, Alzahrani F, et al. Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. Cell reports. 2015; 10(2):148–61. [PubMed: 25558065]

37. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, et al. Exome sequencing identifies the cause of a mendelian disorder. Nature genetics. 2010; 42(1):30–5. [PubMed: 19915526]

38. Kaplan JC, Hamroun D. The 2015 version of the gene table of monogenic neuromuscular disorders (nuclear genome). Neuromuscular disorders: NMD. 2014; 24(12):1123–53. [PubMed: 25625149]

39*. Comprehensive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. Genome biology. 2015; 16:134. A comprehensive review on panels and exome sequencing. [PubMed: 26112015]

40. Guerreiro R, Bras J, Hardy J, Singleton A. Next generation sequencing techniques in neurological diseases: redefining clinical and molecular associations. Human molecular genetics. 2014; 23(R1):R47–53. [PubMed: 24794858]

41. Pasaniuc B, Rohland N, McLaren PJ, Garimella K, Zaitlen N, Li H, et al. Extremely low-coverage sequencing and imputation increases power for genome-wide association studies. Nature genetics. 2012; 44(6):631–5. [PubMed: 22610117]

42. Vasta V, Merritt JL 2nd, Saneto RP, Hahn SH. Next-generation sequencing for mitochondrial diseases: a wide diagnostic spectrum. Pediatrics international: official journal of the Japan Pediatric Society. 2012; 54(5):585–601. [PubMed: 22494076]
43. Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, Beeson KY, et al. Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome biology. 2009; 10(3):R32. [PubMed: 19327155]
44. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. Nature reviews Genetics. 2011; 12(11):745–55.
45. Bell CJ, Dinwiddie DL, Miller NA, Hateley SL, Ganusova EE, Mudge J, et al. Carrier testing for severe childhood recessive diseases by next-generation sequencing. Science translational medicine. 2011; 3(65):65ra4.
46. Goldstein DB, Allen A, Keebler J, Margulies EH, Petrou S, Petrovski S, et al. Sequencing studies in human genetics: design and interpretation. Nature reviews Genetics. 2013; 14(7):460–70.
47. Kiezun A, Garimella K, Do R, Stitziel NO, Neale BM, McLaren PJ, et al. Exome sequencing and the genetic basis of complex traits. Nature genetics. 2012; 44(6):623–30. [PubMed: 22641211]
48. Schuster SC. Next-generation sequencing transforms today's biology. Nature methods. 2008; 5(1): 16–8. [PubMed: 18165802]
49. Houlden H, King RH, Hashemi-Nejad A, Wood NW, Mathias CJ, Reilly M, et al. A novel TRK A (NTRK1) mutation associated with hereditary sensory and autonomic neuropathy type V. Ann Neurol. 2001; 49(4):521–5. [PubMed: 11310631]
50. Sailer A, Scholz SW, Gibbs JR, Tucci A, Johnson JO, Wood NW, et al. Exome sequencing in an SCA14 family demonstrates its utility in diagnosing heterogeneous diseases. Neurology. 2012; 79(2):127–31. [PubMed: 22675081]
51. Singleton AB, Hardy J, Traynor BJ, Houlden H. Towards a complete resolution of the genetic architecture of disease. Trends Genet. 2010; 26(10):438–42. [PubMed: 20813421]




Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

**KEY POINTS**

- Neuromuscular disorders are clinically heterogeneous. There is a broad spectrum of disease ranging from myopathy and CMT through to the more aggressive neuromuscular impairment in MND.
- Next-generation sequencing technologies have helped greatly in diagnosing these disorders and underpinning the genetic pathways.
- Technology has evolved to create large disease gene panels through to whole exome sequencing, and final transition to whole genome sequencing.
- The understanding of genetic causes and disease pathways in the remainder of neuromuscular disorders will be key to identifying treatments.
- The formation of collaborative groups such as the neurology and neuromuscular GeCIP will greatly advance the interpretation and functional understanding of whole genome sequencing.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

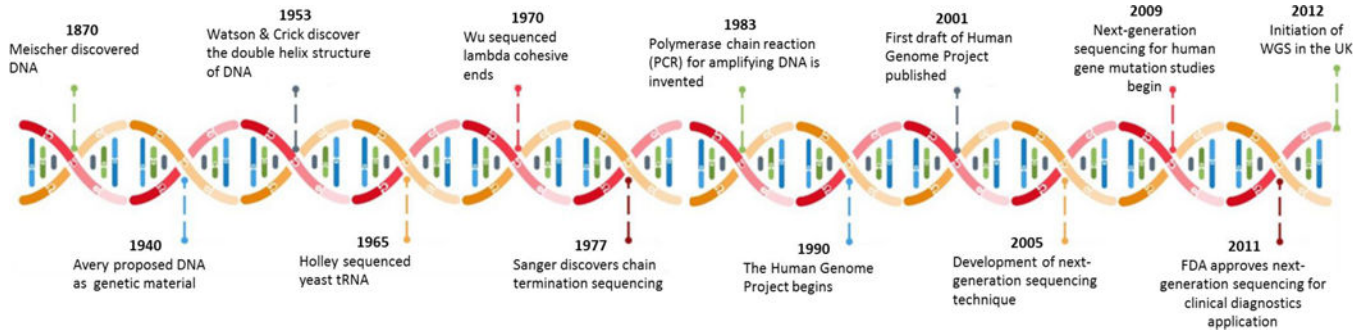


**Figure 1A. A timeline depicting the key events in the history of genomics**

Genetic research and the involvement of new technologies have played a major role in the increase of the amount of DNA sequence generated per person per year at a greater cost efficiency. From the discovery of DNA in 1870, to the introduction of chain terminators to DNA sequencing and separation of DNA tracts by polyacrylamide gel electrophoresis (PAGE) through to the implementation of whole exome (WES) and whole genome sequencing (WGS) in diagnostic and clinical research. Without the integration of the DNA sequencing chemistry and DNA cloning, the sequencing of whole genomes would not be possible today. The innovation of laser detection systems and the progress in computational hardware and analytical software have greatly accelerated the speed of DNA sequencing, as well as its cost and labour needs.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

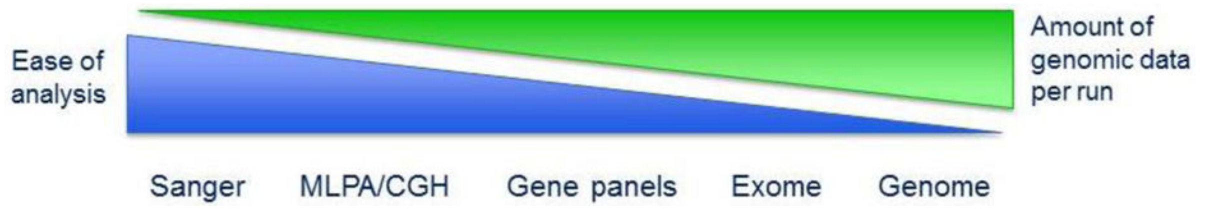


**Figure 1B.**

Schematic representation comparing the ease of analysis compared to the amount of genomic data generated per run.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

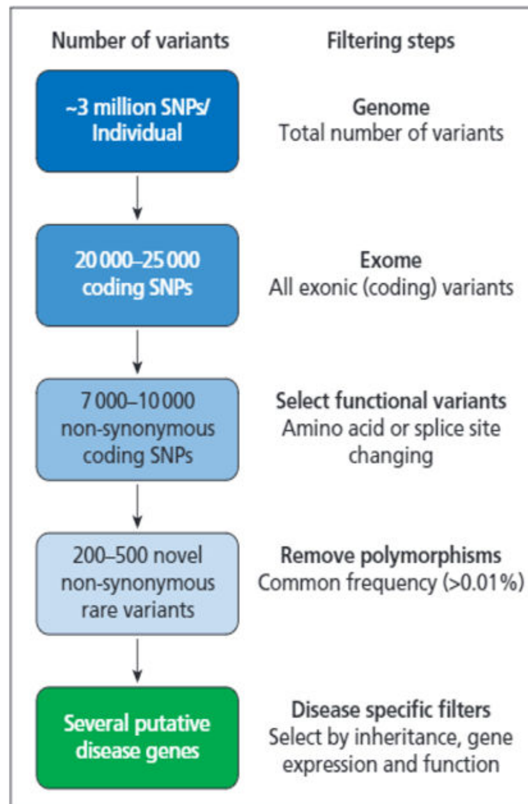


**Figure 2. Filtering and functional analysis technique for whole exome sequencing**
Data analysis of exomes involves the alignment of sequences, removing duplicates and annotating the data sequence and variants. When investigating disease-causing mutations, common polymorphisms and non-coding amino acid changes are removed to leave a list of unique or very rare (<0.01% of the population) variants that are either heterozygous or homozygous.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

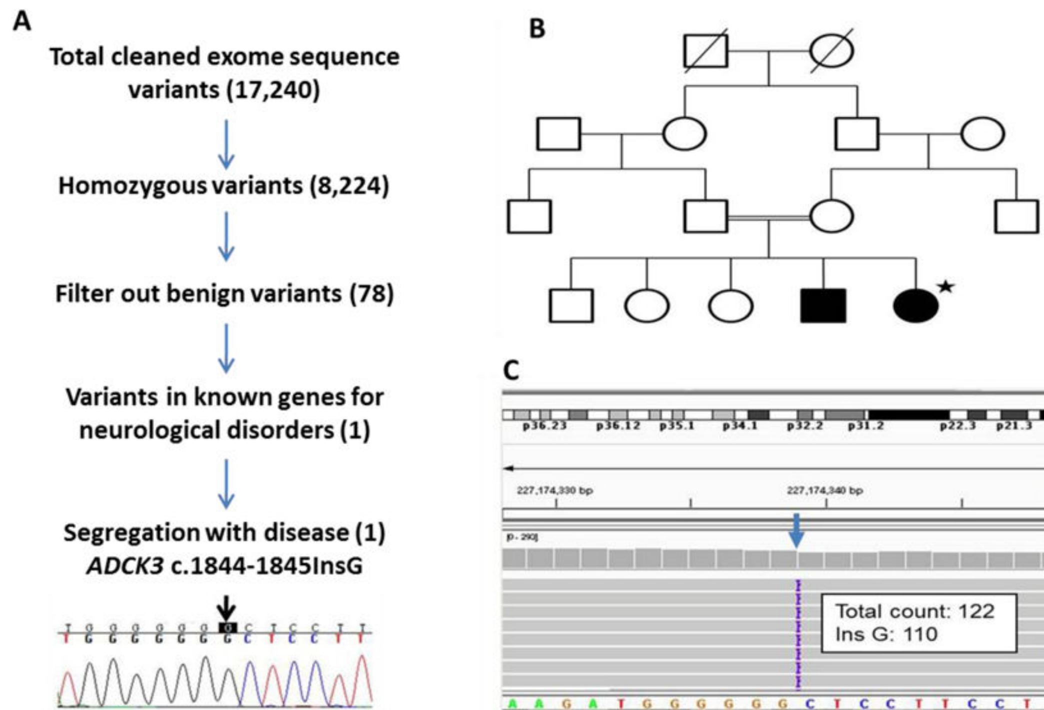


**Figure 3. An example of exome sequencing filtering strategy in a recessive family**

The analysis of multiple family members is often very important in proving a genetic variant is pathogenic. A. Exome filtering of variants. B. Family tree and C. Exome Viewer visualising the mutation.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



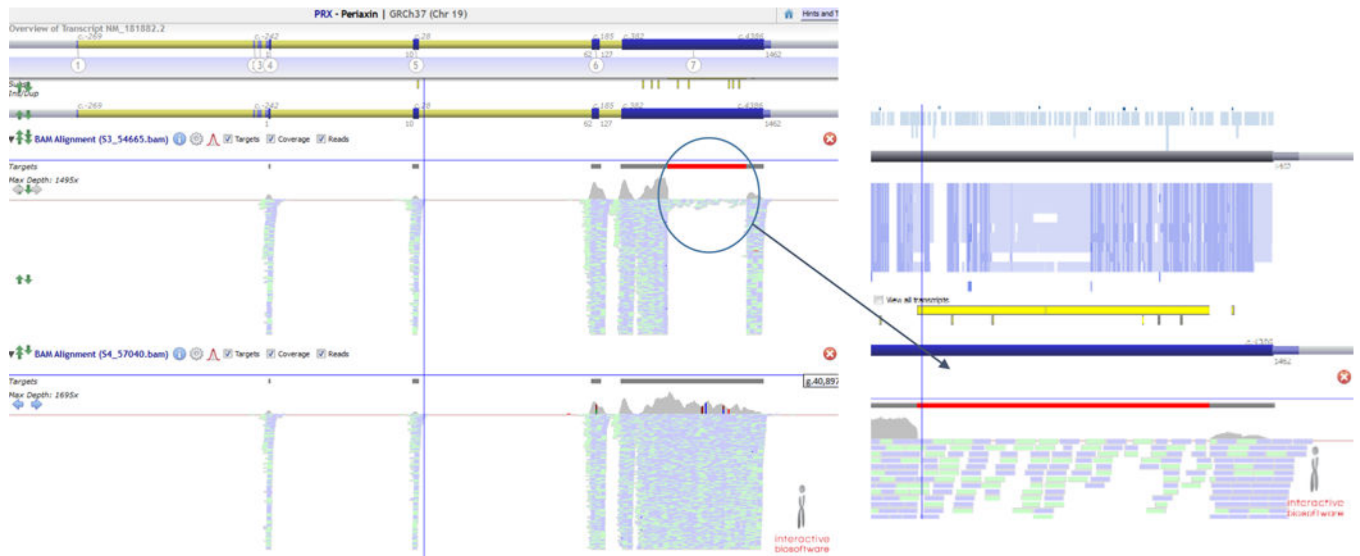


**Figure 4. A deletion in exon 7 of the (Periaxin) *PRX* gene, investigated by a gene panel and visualized with an interactive software**

In the top panel, the patient has part of the gene deleted (indicated by a circle) which is determined by the greatly reduced coverage depth of that region (red). The bottom panel shows a healthy control patient which has a uniform coverage depth throughout exon 7 of *PRX* gene.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

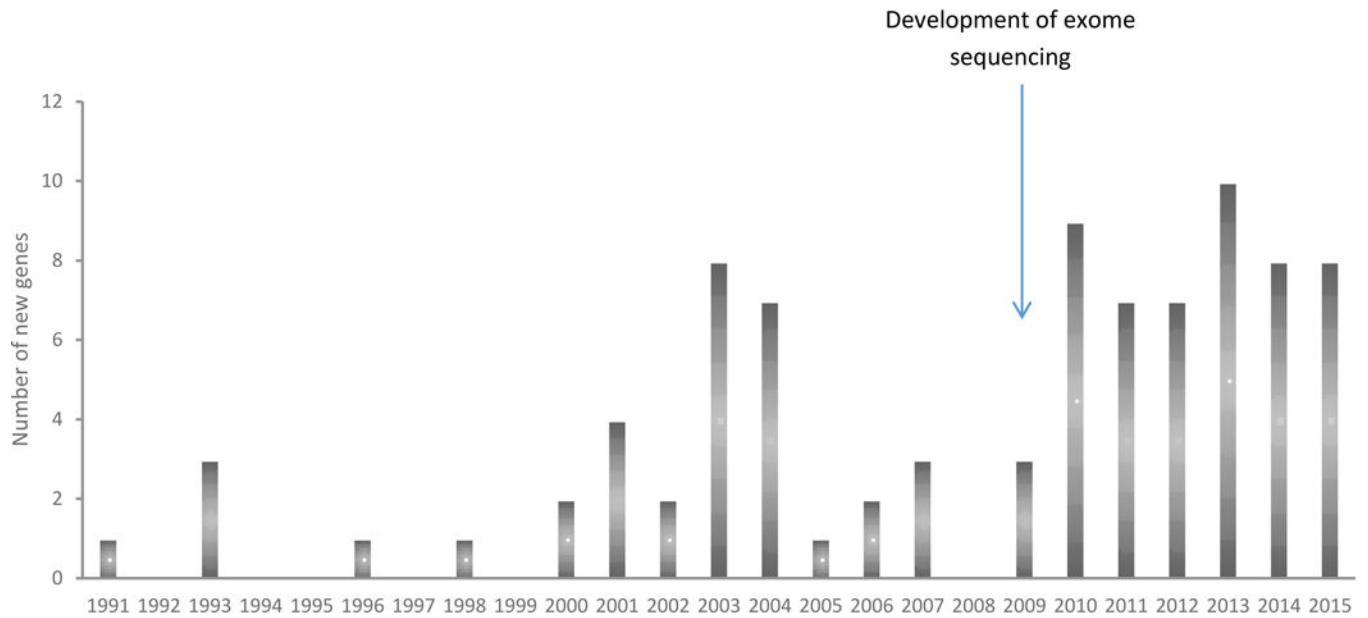


**Figure 5.**

A timeline of discovery of genes involved in Charcot-Marie-Tooth disease. In the early 2000's linkage analysis and Sanger sequencing is widespread but requires larger families. After 2009/2010 the advent of next generation sequencing has allowed new disease genes to be identified in smaller and rarer families. X axis = number of genes, Y axis = year.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

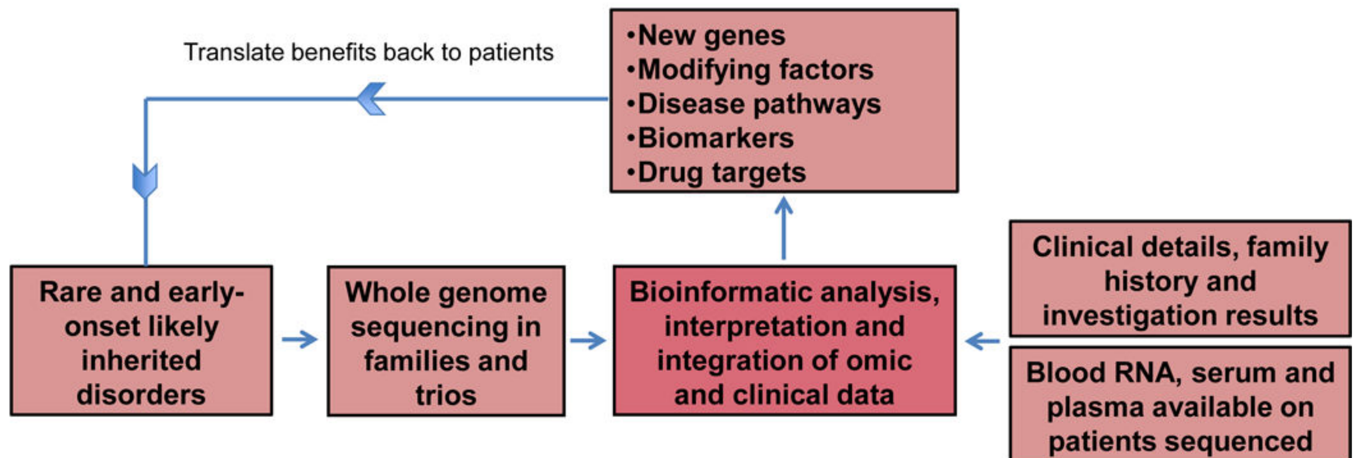


**Figure 6. Genomics England and Genomics England Clinical Interpretation Partnership (GeCIPs) flow diagram**

The major objectives of the Neurology and Neurodegeneration research plan are to advance and optimise whole genome sequencing validation, interpretation and interrogation. This will significantly accelerate our understanding of disease pathophysiology and key mechanistic pathways enabling the future development of new therapies.

**Specific aims include:**

1. Development of data capture and similarity scoring algorithms to delineate rare and severe neurological and neurodegenerative disorders into homogeneous phenotypic groups.
2. Identification of novel disease genes, genetic risks and modifying factors to enable comprehensive clinical diagnostic testing, prediction of disease onset and severity.
3. Gene discovery will lead to the identification of novel pathways that will be investigated and modeled to advance our understanding of disease pathophysiology and mechanisms.
4. Training to develop the next generation of NHS technologists, scientists and clinicians in genomic medicine to sustain a thriving effective team for the future.
5. Future collaboration with industry to utilise pathway discovery and identify novel series of medicines, vaccines, and pharmacogenomics to deliver precision patient treatments based on genomics.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

**Table 1**

Current techniques used to diagnose neuromuscular conditions in general and specific defects in NMD and how NGS will start to change these methods. With longer read lengths WGS will be able to pick up more repeat expansions. WGS also includes mitochondrial genome probes where the mitochondrial genome is usually sequenced at >1000 fold coverage.

| Neuromuscular Disease | Genetic defect | Inheritance | Current method | Next generation sequencing method |
| --- | --- | --- | --- | --- |
| Majority of inherited neuromuscular conditions | Single gene point mutations and small deletions or duplications | AD, AR and X-linked | Sanger sequencing and gene panels | NGS panels and WES, moving to WGS in the future |
| Inherited neuromuscular conditions | Single exon or larger deletions or insertions in neuromuscular genes | AD, AR and X-linked | CGH for screening or MLPA when targeted | WGS, this will pick up point mutations and deletions/ insertions |
| Charcot-Marie Tooth disease type 1A | Chromosome 17p11.2 duplication | AD | MLPA | CGH for ease and speed but also WGS |
| HNPP | Chromosome 17p11.2 deletion | AD | MLPA | CGH for ease and speed but also WGS |
| Duchene Muscular Dystrophy | Large re-arrangements and single exon deletions | X-linked | MLPA or CGH array | CGH for ease and speed but also WGS |
| Oculopharyngeal muscular dystrophy | Polyalanine repeat small expansions | AD/AR | Fragment analysis | WGS, this will pick up point mutations and deletions/ insertions |
| MD1 | CTG repeat expansions | AD | Fragment analysis | Fragment analysis |
| MD2 | CCTG repeat expansions | AD | Fragment analysis | Fragment analysis |
| Kennedy's | CAG repeat expansions | X-linked | Fragment analysis | Fragment analysis |
| MND – C9orf72 | GGGGCC repeat expansions | AD | Fragment analysis | Fragment analysis |
| Mitochondrial (ideally in DNA from muscle tissue) | Point mutations, single and multiple deletions | Maternal | Southern blot, long PCR or Sanger sequencing | Targeted mitochondrial sequencing or WGS that includes mitochondrial |

Key: CGH = cytogenetic genomic hybridization. WES = whole exome sequencing. WGS = whole genome sequencing. MND = motor neurone disease. AD = autosomal dominant. AR = autosomal recessive. MD = myotonic dystrophy. HNPP = Hereditary neuropathy with pressure palsies. MLPA = multiple ligation probe analysis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

**Table 2**

Pros and cons of NGS techniques: exome sequencing, genome sequencing, Sanger sequencing, fragment analysis, MLPA/CGH arrays and gene panels.

| **Technique** | **Sanger sequencing** | **Fragment analysis** | **MLPA/CGH array** | **Gene Panels** | **Exome sequencing** | **Genome Sequencing** |
|---|---|---|---|---|---|---|
| **Detectable variants** | Point mutations and small deletions or insertions | Repetitive sequences such as expanded CAG repeats | MLPA: small exon level copy number changes<br>CGH: small to large copy number changes | Point mutations and small to whole gene deletions or insertions | Point mutations and small deletions or insertions | All variants, including large deletions or insertions |
| **Regions** | Single region per run usually up to 500bp but can go up to 1000bp with certain DNA templates. | Single genetic region | MLPA: specified gene regions.<br>CGH: varies from low to very high genome to density | Targeted genomics regions such as all exons in specific genes, up to several thousand genes | Exons and flanking introns in the genome | Genome but not long repetitive regions (>150bp) |
| **Coverage and error rates** | Very accurate, >99.9% when visualised | Very accurate, >99.9% when visualised | Very accurate,<br>Regions visualised | High coverage, aimed at 99.5% and >50X | High coverage, aimed at 96% at >50X.<br>GC regions such as exon 1 of genes often poorly covered | Genome but not long repetitive regions (>150bp) |
| **Starting material** | 25ng of DNA | 25ng of genomic DNA | 100ng of genomic DNA | 500ng of genomic DNA | 1μg of genomic DNA | 1μg of genomic DNA |
| **Data analysis** | Pattern recognition and requires training and practice. Sanger sequencing is ideally done in forward and reverse. | | | Size of data is challenging, detailed pipeline required and requires expert bioinformatics | | |
| **Cost** | Single run £10 but expensive per base | Single run £10 | £100–500 per run depending on size of array | Around £140 per run | Around £280 per run | Around £800 per run |