

New generation  
sequencing  
technologies:  
population genetics



**ASSIGN  
BUSTER**

New  
generation  
sequencing  
technologies:  
implications  
for the  
science of  
population  
genetics

- Kate  
Elizabet  
h  
Moffitt

New generation sequencing technologies have the potential to rapidly accelerate population genetics research, allowing scientists to comprehensively understand complex evolutionary histories, as well as functional and ecological biodiversity (Shokralla, *et al.* , 2012; Shendure & Hanlee, 2008). Prior to 1977, sequence production involved the handling of toxic chemicals and radio-active isotopes, restricting sequencing to persons of high expertise and speciality (Hunkapiller, 1991; Swerdlow, *et al.* , 1990; Sanger, *et al.* , 1977). In 1977, Fred Sanger and Alan R. Coulson published two methodological papers describing a new form of DNA sequencing technology, which would lead to the method (capillary-based, semi-automated Sanger biochemistry) used almost exclusively in the field, for the

next 30 years (Shendure & Hanlee, 2008). Sanger sequencing transformed biology. It became a tool for deciphering complete genes and, later, entire genomes. Due to the unprecedented extent at which Sanger technology grew, factory-like enterprises, called sequencing centres, were established, housing hundreds of DNA sequencing instruments, operated by cohorts of personnel (Schuster, 2008; Hunkapilla, *et al.*, 1991).

Despite the dominance of Sanger sequencing in laboratories, for a number of decades, the technology had and continues to be hampered by inherent limitations in throughput, scalability, speed and resolution (Shendure & Hanlee, 2008). To overcome these barriers, an entirely new technology was required, one that democratised the field, putting the technology of comprehensive genetic analysis into the hands of individual investigators, not only major genome research centres (Shendure & Hanlee, 2008). The need for new technologies was pushed for by the facilitators of the Human Genome Project (HGP) (Ventor, *et al.*, 2001). The excitement and successful completion of the HGP, by two competing research bodies, led to collective hunger for more advanced, economical sequencing technologies. Next-generation sequencing (NGS), also known as massively parallel sequencing, was such a technology and has ignited a revolution in genomic science, similar to that seen when Sanger technology was presented in 1977, honing in on the era of 'post-genomic' research (Schuster, 2008).

The revolutionary nature of NGS technologies first became apparent in 2005, in two separate publications, *454 Life Sciences* (Marguiles, *et al.*, 2005) and the Multiplex Polony Sequencing Protocol (Shendure, *et al.*, 2005). The methodology of both research groups resulted in vast reductions in the

<https://assignbuster.com/new-generation-sequencing-technologies-population-genetics/>

necessary reaction volume, while dramatically extending the number of sequencing reactions (Schuster, 2008). Despite such advances, in sequencing technology, NGS had a slow uptake in the scientific community, with a number of scientists having reservations. According to Schuster (2008), scientists accustomed to Sanger sequencing, as well as the initial scepticism echoed by funding bodies, resulted in a fear that large financial investments into Sanger-sequencing technologies would not produce returns, due to the technologies becoming obsolete. Other concerns were also raised, regarding the sequencing fidelity, read length, infrastructure cost and the handling of the large data volumes produced by NGS (Zhang, *et al.*, 2011). It was the process of combining ongoing Sanger sequencing projects with NGS technologies that promoted its acceptance, into the scientific community. Once the enormous potential of the technology had been realised, along with new and upcoming biology projects that required sequencing outside of what the current Sanger technology could feasibly produce, the concerns raised by NGS's early sceptics started to be overlooked. A combination of both first and second generation technologies are now used in sequencing facilities and projects around the world, the implications of which, for the fields of evolutionary biology and population genetics is vast.

Researchers now have the ability to observe small changes in ecological community structure that may occur following anthropogenic or natural environmental fluctuations (Hajibabaei, *et al.*, 2011; Leininger, *et al.*, 2006; Hunkapiller, *et al.*, 1991). Such implications of NGS technologies has led to the generation of whole-genome sequence data, for thousands of individuals

(Akey & Shriver, 2011; Harismendy, *et al.*, 2009). The availability of such data is leading to a better understanding of evolutionary processes, such as descriptions of sex-biased dispersal and mutation rate biases (e. g., Wilson Sayres, M. A., *et al.*, 2011). Furthermore, the ability to sequence the genomes of species, that have been long extinct, is no longer nonsensical, provided the samples from which DNA is to be extracted is still viable (Green, *et al.*, 2010; Reich, *et al.*, 2010). The hope that such projects may help population geneticists better understand the process of extinction, whether anthropomorphically or naturally induced, may help those endangered species whose current possibility of extinction, in the near future, is high (Akey & Shriver, 2011; Miller, *et al.*, 2011). However, despite such ambitious aspirations of population geneticists, one large area of research that remains surprisingly unanswered, within the literature, is the definition of a population or ‘the population concept’ (Waples & Gaggiotti, 2006).

Given the importance of such a concept, one might expect to find a commonly used definition, one that is applicable to wild species, to determine how many populations exist within a delineated geographic area and the relationships amongst them (Waples & Gaggiotti, 2006). However, one does not exist, rather there is evidence that what makes a ‘population’ is based on the research question. NGS technologies are providing population geneticists with the opportunity to flesh out a detailed definition of a population, on the molecular level. For example, Waples & Gaggiotti (2006) ask “How different must molecular units be before individuals can be considered a part of separate populations?” Different criteria can be

established and assigned to individuals, in order to determine the answer.

The interplay of different evolutionary forces (selection, migration, drift) will favour different species, with different forces being more obvious, at the molecular level, than others. The ability to pose a research question, pertaining to the individuals, within a particular habitat, is now possible due to the ability to sequence numerous samples with NGS technologies.

The implications, in population genetics, for a new generation of sequencing technologies, are a greater focus on testing expectations. Such expectations, simultaneously, result in excitement and daunt to those undertaking evolutionary and population genetic research, at present. Excitement exists because fundamental questions, pertaining to the patterns of genetic variation, within and between species, can now be analysed, with new generation sequencing technologies, such as NGS. Although NGS technology may still be in its infancy, the powerful possibility of analysing massive data sets is within reach of the individual and large-scale sequencing facilities alike, at a highly reduced cost. However, the methodological tools and theoretical models needed to interpret such large data sets are equally daunting to both new, and experienced, evolutionary and population geneticists. Despite such present and future challenges, population genetics research is looking promising, thanks to advances in NGS adoption and computation.

## References

Akey, J. M. & Shriver, M. D. (2011). A grand challenge in evolutionary population genetics: new paradigms for exploring the past and charting the future in the post-genomic era. *Frontiers in Genetics* 2, 1-2.  
<https://assignbuster.com/new-generation-sequencing-technologies-population-genetics/>

Green R. E., Krause J., Briggs A. W., Maricic T., Stenzel U., Kircher M., Patterson N., ... Pääbo S. (2010). A draft sequence of the Neanderthal genome. *Science* 328, 710–722.

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A. C. & Baird, D. J. (2011). Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS ONE* 6, e17497.

Harismendy, O., Ng, P. C., Strausberg, R. L., Wang, X., Stockwell, T. B., Beeson, K. Y., Schork, N. J., ... Frazer, K. A. (2009). Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biology* 10(3), 32-39.

Hunkapiller, M. W. (1991). Advances in DNA sequencing technology. *Current Opinion in Genetics & Development* 1(1), 88-92.

Hunkapiller, T., Kaiser, R. J., Koop, B. F. & Hood, L. (1991). Large-scale and automated DNA sequence determination. *Science* 254, 59-67.

Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G. W., Prosser, J. I., Schuster, S. C. & Schleper, C. (2006). Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* 442, 806-809.

Marguiles, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., ... Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376-380.

Miller W., Hayes V. M., Ratan A., Petersen D. C., Wittekindt N. E., Miller J., Walenz B., ... Schuster S. C. (2011). Genetic diversity and population

<https://assignbuster.com/new-generation-sequencing-technologies-population-genetics/>

structure of the endangered marsupial *Sarcophilus harrisii* (Tasmanian devil).  
*Proc. Natl. Acad. Sci. U. S. A.* 108 (30), 12348-12353.

Reich D., Green R. E., Kircher M., Krause J., Patterson N., Durand E. Y., Viola B., ... Pääbo S. (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468, 1053–1060.

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A. III, Slocombe, P. M. & Smith, M. (1977). Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* 265, 687-695.

Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nature Methods* 5(1), 16-18.

Shendure, J. & Hanlee, J. (2008). Next-generation DNA sequencing. *Nature Biotechnology* 26(10), 1135-1145.

Shendure, J., Porreca, G. J., Reppas, N. B., Lin, X., McCutcheon, J. P., Rosenbaum, A. M., Wang, M. D., Zhang, K., Mitra, R. D. & Church, G. M. (2005). Accurate multiplex polony sequencing of an evolved bacterial genome. *Science* 309, 1728-1732.

Shokralla, S., Spall, J. L., Gibson, J. F. & Hajibabaei, M. (2012). Next generation sequencing technologies for environmental DNA research. *Molecular Ecology* 21, 1794-1805.

Swerdlow, H., Wu, S. L., Harke, H. & Dovichi, N. J. (1990). Capillary gel electrophoresis for DNA sequencing. Laser-induced fluorescence detection with the sheath flow cuvette. *Journal of Chromatography* 516, 61-67.

<https://assignbuster.com/new-generation-sequencing-technologies-population-genetics/>



Ventor, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Amanatides, P., ..., Zhu, X. (2001). The sequence of the human genome. *Science* 291, 1304-1351.

Waples, R. S. & Gaggiotti, O. (2006). INVITED REVIEW: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology* 15 (6), 1419-1439.

Wilson Sayres M. A., Venditti C., Pagel M., Makova K. D. (2011). Do variations in substitution rates and male mutation bias correlate with life history traits? A study of 32 mammalian genomes. *Evolution* 65 (10), 2800-2815.

Zhang, J., Chiodini, R., Badr, A. & Zhang, G. (2011). The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics* 38, 95-109.