**\section{Next-generation sequencing revolution}**

\textit{Any sufficiently advanced technology is indistinguishable from magic}

\begin{flushright} [Arthur C. Clarke] \end{flushright}

The 20th century had been an amazing decade of series of discoveries in biological sciences. Erwin Schrodinger proposed in his famous lecture “What is Life” in 1943 that physical material responsible for genetic inheritance must have an aperiodic crystal structure. Since the identification of DNA as the transforming agent of R strain to S strain by Oswald Avery \cite{} and the discovery of the structure of the DNA by James Watson and Francis Crick in 1954 \cite{}, there is now no question that DNA is the physical material responsible for genetic inheritance and the genome, the entire corpus of genetic information of an organism, dictates the embryonic development, cell differentiation and function, organisation of tissues, and even the life span of an organism.

Today, sequencing instruments function as a modern microscope not to look at physiology of individual or population of cells, but to measure and interpret genetic variations (DNA) and their phenotypic manifestations (RNA and amino acids) to understand the intricacies of cellular functions. Ion-torrent \cite{}, Pyro \cite{}, Roche 454 \cite{} and Illumina \cite{} sequencing platforms once competed in the sequencing market as next-generation sequencing platforms and Illumina platform has emerged as the sequencing method of choice such that Illumina platform has become synonymous with the next-generation sequencing. The advent of high-throughput sequencing instrument combined with exponential decrease in sequencing costs has completely transformed how we approach biological questions.

**\subsection{Illumina Platform}**

Illumina platform performs sequencing-by-synthesis (SBS) to generate paired-end reads, each of 150bp read length. A typical read from the Illumina platform is shorter than that produced from Sanger sequencing, but the shorter read length is compensated with redundant sequencing of the genome.

***contact, detect, dissociate, de-block, incorporate, rinse, repea***

%% sequencing cost, moore’s law

%% economics of scale

%% innovation on top of next-generation sequencing

%% ancient genomics

%% iteration of illumine sequencing machines: GAII

%% hi-c sequencing: loops configurations, topologically associating domain (TADs), A/B compartments

%% chip sequencing

%% 3C sequencing (Job Dekker)

The technical limitations of Illumina sequencing (base accuracy and short read length), however, has been the bottleneck for improving rare genetic disease diagnostics yield, detecting rare somatic mutations and constructing high-quality reference genomes for non-human species. De novo assembly of other species, previously, have been attempted using de brugjin graph based de novo assembly algorithms with short reads, but assemblies produced from short reads were highly fragmented and incomplete. In addition, scaffolding strategies often did not provide sufficient long-range information to produce chromosome-level pseudomolecules and as a result, these assemblies provided incomplete information for comparative genomics purposes. Hence, assemblies produced from short reads often have collapsed repeats or contigs that cannot be placed accurately.

%% population-genetics

%% population genomics

%% increase in the number of whole-genome sequenced with illumine sequencing

%% cancer genomics, driver mutation, mutational signatures

%% clinical sequencing

%% tumour evolution

%% liquid biopsy

The human genome project is estimated to have cost 3 billion dollars, equivalent to 1 dollar per base pair. And as technology becomes more ubiquitous and democratised, we have constantly shifted/moved from studying one individual to studying the group. We initially focused on studying a single individual and as sequencing cost has decreased, population genomic studies and the history of differences and going back in time to study our lineage. What is common and different.

**\subsection{Data Standardisation}**

A wave of standardisation to create file formats that is universally accepted across the community.

The technical limitations of the Illumina platform limits the interrogation of the genome and the inability to access the “dark matter” of the genome \cite{} and to improve clinical diagnosis with Illumina platform has convinced other researchers to use other sequencing technologies.

**\subsection{Short-read sequencing applications}**

**\subsubsection{Bisulfite sequencing}**

**\subsubsection{Single-cell sequencing}**

**\subsubsection{High-throughput chromatin conformation capture sequencing}**

**\subsubsection{}**