Advocacy – How to bring WGS to the attention of decision makers at the Institute

Antonio Battisti, DVM, Alessia Franco, DVM
Istituto Zooprofilattico Sperimentale del Lazio e Toscana,
National Reference Laboratory for Antimicrobial Resistance, Rome, Italy

The use of the Whole Genome Sequencing (WGS) in Monitoring of Antimicrobial Resistance
EURL-Training Course 2017
DTU, Kgs. Lyngby 27th September 2017

Advocacy (definition)

- (Public) support for or recommendation of a particular cause or policy

Synonyms:

support for, **argument for**, arguing for, calling for, **pushing for, pressing for…**

Outline of a «systematic approach» to the purchase proposal to decision makers

1. Introduction
2. Reasons for purchasing by the Department, NRL-AR, Italy
3. Logistics
4. Comparison of technologies and instruments on the market
4.a. Comparison of at least two “instruments” A and B (considering a market segment of small-medium throughput)

4.b. comparison of protocols preliminary to sequencing, **and “technology-dependent”**
4.c. **analysis of the pros and cons of the technical characteristics** and **accuracy of the technologies to be compared (in relation to the main use at lab);**
5. Other instruments required during steps other than “sequencing” (see 4.b),

6. Options for purchase: plan and expected costs.
7. Costs of reagents: average sequencing costs per genome (including personnel costs per genome sequenced!!!)
8. Sharing the technology within the Institute: possible use by other departments and possible scenario and evolution for institutional purposes.

Prerequisites («protective variable…»)

Human factor:

- Try to «get in tune» with the General Director…

Ingredients:

-Reputation, credibility from your side…

-Wisdom and farsightedness on the General Director’s side

Usually both are of some help….

Next Generation Sequencing: an Intro for our General Director (& Managers)

DNA Sequencing (from Wikipedia)

- **DNA sequencing** is the process of determining the precise order of nucleotides within a DNA molecule. It includes any method or technology that is used to determine the order of the four bases—adenine, guanine, cytosine, and thymine—in a strand of DNA. The advent of rapid DNA sequencing methods has greatly accelerated biological and medical research and discovery.[1]
- Knowledge of DNA sequences has become indispensable for basic biological research, and in numerous applied fields such as medical diagnosis, biotechnology, forensic biology, virology and biological systematics. The rapid speed of sequencing attained with modern DNA sequencing technology has been instrumental in the sequencing of complete DNA sequences, or genomes of numerous types and species of life, including the human genome and other complete DNA sequences of many animal, plant, and microbial species.

NGS (from literature and current knowledge…)

- Next-generation sequencing (NGS) utilizes massively parallel sequencing to generate thousands of megabases of **sequence information** per day. Next-generation techniques are based on a "sequencing by synthesis" principle, where nucleotides incorporated into a strand of DNA provide a unique signal, or a “pH change” principle, where the signal is in the form of a pH change.
- **As in other medical fields**, the availability of next generation sequencing (NGS) techniques has revolutionized both **diagnostics and surveillance** of **infectious diseases and AMR**.
- **NGS is able to replace**, for certain purposes, the current **“reference method”** (Sanger Sequencing).
- Additionally, it is capable to provide Sequencing of the Whole Genome of pathogenic microrganisms (e. g. bacteria, viruses, protists) at sustainable costs, for surveillance and molecular epidemiology purposes, **at multisite level**.
- The use of NGS/WGS takes advantage of the **increasing availability and speed and decreasing cost** per base of NGS offered by deep sequencing machines.
- **“These aspects have lead to a democratization of genomics”**: Whole Genome Sequencing at affordable costs for (Veterinary) Public Health Laboratories

Reasons for purchasing at Department level

- The instrument is needed to **facilitate the mission (Accredited multisite laboratory following ISO 17025 rules) of the Department** (Institute) and allow:

- **deep molecular characterization** (ID, genetic basis of virulence **and AMR**, population structure, phylogeny etc.);

- **molecular epidemiology investigations** (tracing back, source attribution etc.) on major pathogens, zoonotic agents, **AMR determinants and their genetic environment**.

- Specifically, it is an instrument that can provide sequences of "whole genomes" of pathogenic agents (e. g. the entire genetic content of a pathogenic agent consisting of millions of nucleotides, known as "Whole Genome Sequencing") **in an accurate, fast, and cost-effective way**.
- **The above described is not feasible through traditional sequencing technology ("Sanger Sequencing“: we already have one machine with 8 capillaries)**. It should also be emphasized that the above-mentioned NGS technology is not currently available at our Institute.
- For a start, and for expected initial volume of activity, a machine with a medium/small output (e. g. min. 10 Gigabytes, with 25 million reads of at least 100-300 pairs of theoretical bases) would be sufficient, considering that the rapid evolution of in this field will provide a fairly fast "generational replacement” of technologies and related equipment.
- **Prudent approach to the initial investment on a new technology (in this case is around 100,000 Euro)**!!!

Center for Genomic Epidemiology

Home | Organization | Project | Services | Contact

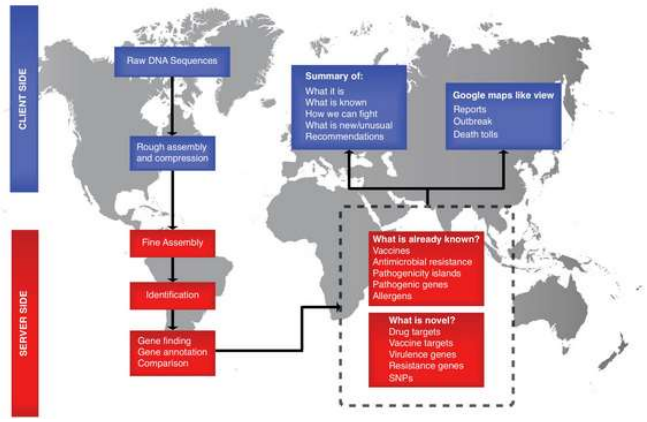
Services

Phenotyping:

- Identifcation of acquired antibiotic resistance genes.
  ResFinder
- Identifcation of functional metagenomic antibiotic resistance determinants.
  ResFinderFG
- Identifcation of acquired antibiotic resistance genes using Kmers.
  KmerResistance
- Prediction of a bacteria's pathogenicity towards human hosts.
  PathogenFinder
- Identifcation of acquired virulence genes.
  VirulenceFinder
- Determination of Restriction-Modification sites (based on REBASE .)
  Restriction-ModificationFinder
- SPIFinder identifies Salmonella Pathogenicity Islands
  SPIFinder

Typing:

- Multi Locus Sequence Typing (MLST) from an assembled genome or from a set of reads
  MLST
- PlasmidFinder identifies plasmids in total or partial sequenced isolates of bacteria.
  PlasmidFinder
- Multi Locus Sequence Typing (MLST) from an assembled plasmid or from a set of reads
  pMLST
- Prediction of bacterial species using a fast K-mer algorithm.
  KmerFinder



Welcome to the Center for Genomic Epidemiology

The cost of sequencing a bacterial genome is $50 and is expected to decrease further in the near future and the equipment needed cost less than $150 000. Thus, within a few years all clinical microbiological laboratories will have a sequencer in use on a daily basis. The price of genome sequencing is already so low that whole genome sequencing will also find worldwide application in human and veterinary practices as well as many other places where bacteria are handled. In Denmark alone this equals more than 1 million isolates annually in 15-20 laboratories and globally up to 1-2 billion isolates per year. The limiting factor will therefore in the future not be the cost of the sequencing, but how to assemble, process and handle the large amount of data in a standardized way that will make the information useful, especially for diagnostic and surveillance.

The aim of this center is to provide the scientific foundation for future internet-based solutions where a central database will enable simplification of total genome sequence information and comparison to all other sequenced including spatial-temporal analysis. We will develop algorithms for rapid analyses of whole genome DNA-sequences, tools for analyses and extraction of information from the sequence data and internet/web-interfaces for using the tools in the global scientific and medical community. The activity is being expanded to also include other microorganisms, such as vira and parasites as well as metagenomic samples.

News

Center for Genomic Epidemiology spinout
June 2016
A spinout company has recently been founded on the basis of Center for Genomic Epidemiology. Read more....

What Can We Learn from a Metagenomic Analysis of a Georgian Bacteriophage Cocktail?
December 2015
Link to article....

WGS typing is a superior alternative to conventional typing strategies
August 2015
In combination with other available WGS typing tools, E. coli serotyping can be performed solely from WGS data, providing faster and cheaper typing than current routine procedures. Link to article....

Introduction to microbial whole genome sequencing and analysis for clinical microbiologist
April 2015
We offer clinical microbiologists the possibility to learn how to use the tools for e.g. typing, identifying plasmids, antibiotic resistance and virulence genes and for phylogenetic analysis.
Sign up....

Consortium to combat infectious disease outbreaks
January 2015
The COMPARE project has been funded with 20 million Euros from the EU. The Consortium consists of 29 partners with multidisciplinary expertise in human health, animal health and food safety. Read more....

Benchmarking of Methods for Genomic Taxonomy
April 2014
How to optimally determine taxonomy from whole genome sequences. Link to article...

CGE tools applied for bacteriophage characterization
March 2014

A key component in the mission of (Veterinary) Public Health Institutes

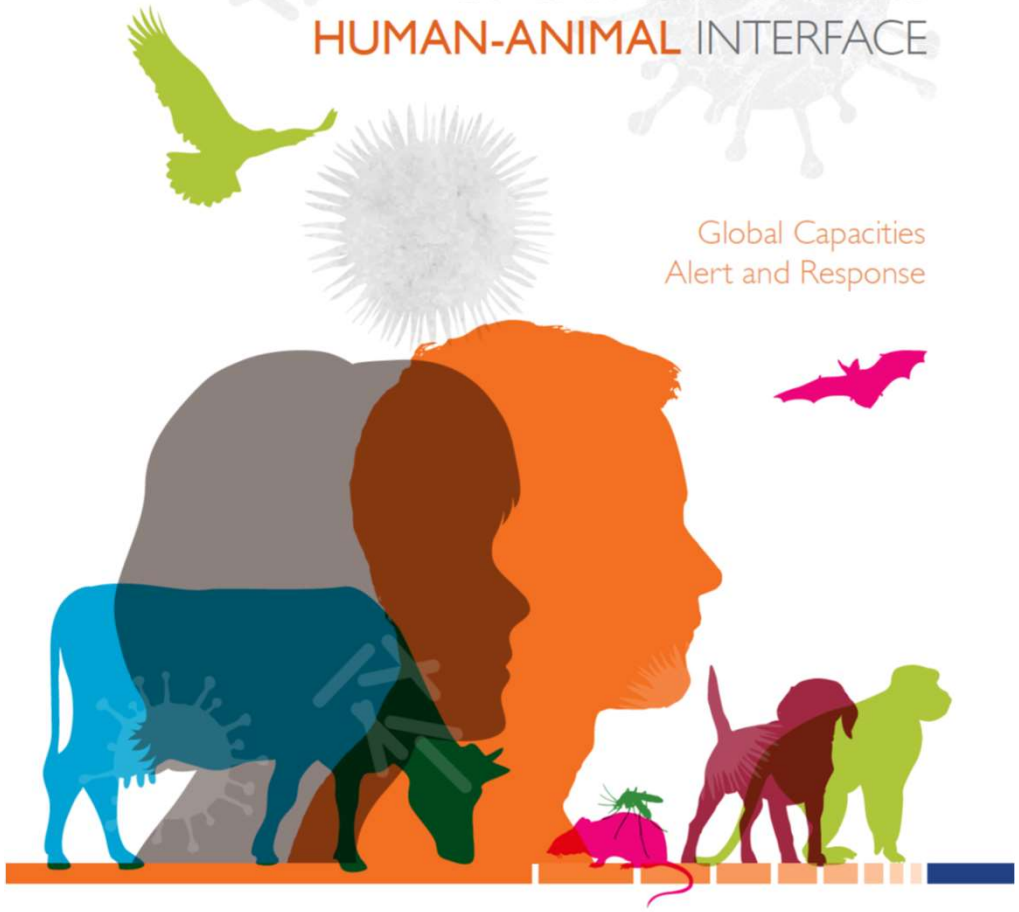
- **Innovation** is an essential component for (Veterinary) Public Health Laboratories when pursuing the objectives of **(Veterinary) Public Health and Consumer Protection ("magic words"…)**

- **Final (rhetorical) Question** to our GD:
- Would you risk to be cut off from this revolution, and innovation ongoing in the "core business" sector of our Institute, or be left behind?
- **Competition** for funding & Research Projects, **beside reputation**, **is based on** "**Diagnostic Capacities**" of Laboratories, to be available/delivered in "real time" (here and now…)

HANDBOOK

FOR THE ASSESSMENT OF CAPACITIES AT THE HUMAN-ANIMAL INTERFACE

Global Capacities Alert and Response



OIE WORLD ORGANISATION FOR ANIMAL HEALTH

World Health Organization

EPIDEMIOLOGY & LABORATORY CAPACITY FOR INFECTIOUS DISEASES

ESSENTIAL FUNDING FOR PUBLIC HEALTH LAB RESPONSE

BACKGROUND:

The Epidemiology and Laboratory Capacity (ELC) cooperative agreement is the US Centers for Disease Control and Prevention (CDC)'s national funding strategy to support state, local and territorial capacities for emerging infectious disease control. In FY2016, ELC funding reached all 50 states, 8 US territories and 6 US cities. Investments strengthen public health workforce, disease detection systems, laboratory capacity and health information capacity. Recently, ELC has been utilized to help with emergency response to emerging infectious diseases such as Zika and Ebola.

Public health laboratories rely on ELC funding in two ways:

1. Flexible funding
   - The Prevention and Public Health Fund (PPHF), established through the Affordable Care Act (ACA), provides $40 million annually to support a highly trained laboratory workforce, develop modern and well-equipped public health laboratories and assist with integrating laboratory and epidemiology functions
   - Health information system infrastructure is supported through ELC and allows for timely exchange of data between public health laboratories and CDC to help make decisions that impact the public's health
2. Categorical or disease-specific funding
   - Emergency funding for Zika and Ebola helped public health laboratories deal with the increased demand in testing by allowing them to purchase necessary laboratory equipment and supplies

   - Allows for rapid testing in outbreak response such as foodborne diseases and influenza
   - Monitors antimicrobial resistance
   - Supports testing capacity for healthcare-associated infections, tickborne diseases, West Nile virus, dengue fever, parasitic diseases and vaccine preventable diseases such as measles, mumps, pertussis and rubella

PREVENTION AND PUBLIC HEALTH FUND SUPPORTS LAB STRENGTHENING

The portion of the ELC's funding that comes from PPHF funds cross-cutting and flexible support for infectious disease epidemiology, laboratory and health information systems.

Investments made by PPHF funds have allowed state and local health departments to strengthen and integrate their capacity to detect and respond to infectious disease and other public health emergencies. For example, ELC funds have made it possible to increase the use of electronic laboratory reporting, improved information technology infrastructures, streamlined program coordination and expanded training activities.

ASSOCIATION OF PUBLIC HEALTH LABORATORIES
8515 GEORGIA AVENUE, SUITE 700, SILVER SPRING, MD 20910 | WWW.APHL.ORG
240.485.2745 (P) | 240.485.2700 (F)

Suppose your General Director says:

«Yes, we want to **invest progressively** on NGS technology and we want to purchase an instrument which is fit for purpose»

«Please, provide us with a thorough analyisis and a **final proposal on what to buy** ,based on current procedures used in Public Health Institutes. The document should be available for any inspection by an external assessment body, according to our national (EU) legislation».

Then, is the Lab responsibility to provide the analysis…

Lab responsibility

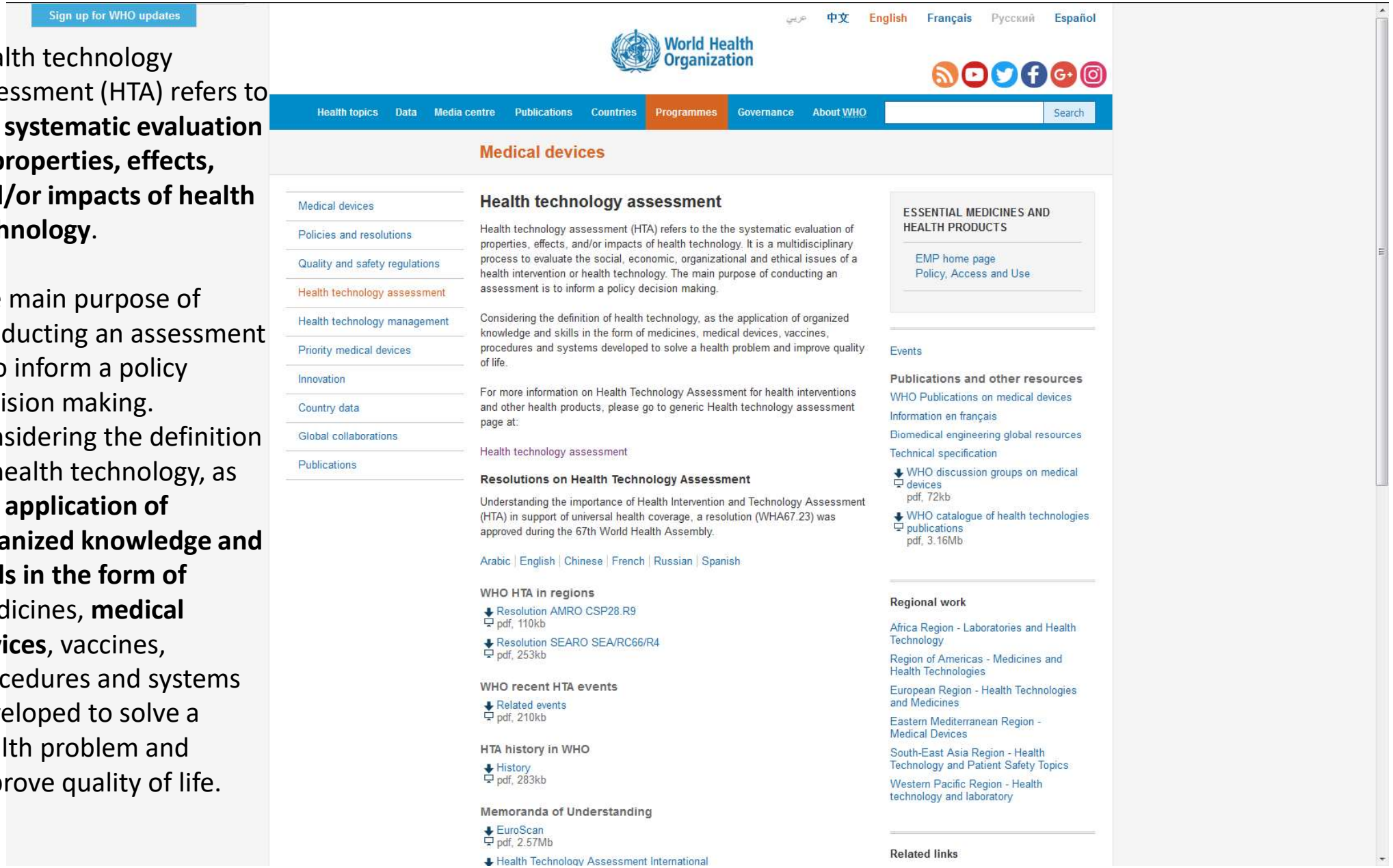
- When there is the «green light» to an investment on NGS:
- It is the Lab responsibility to provide **a report on advantages and impacts of this «new technology»** and propose the purchase of the technology/machine **most suitable for the «core business» of the Lab/Department/Institute.**
- Cancer Research? Rare Diseases Genomics? **Infectious Diseases and AMR**?
- The scope for the use is fundamental in the choice, considering technologies/instruments available on the market
- Science-based, quantitative parameters available in the literature usually help in taking a decision when comparing technologies: for «brand new technologies it may be more difficult…»
- **In the case of NGS technologies/machines: Accuracy, Error Rate…**

Health technology assessment (HTA) refers to the **systematic evaluation of properties, effects, and/or impacts of health technology**.

…

The main purpose of conducting an assessment is to inform a policy decision making.

Considering the definition of health technology, as **the application of organized knowledge and skills in the form of** medicines, **medical devices**, vaccines, procedures and systems developed to solve a health problem and improve quality of life.



Some practical experience on how to assess the technology and propose what to buy

4. Comparison of **technologies** and instruments on the market
4.a. Comparison **of at least two** "**instruments**" A and B (considering a market segment of small-medium throughput) based on different technologies

4.b. comparison of **protocols preliminary to sequencing and asses which are "technology-dependent" and require more labour or additional instruments**

4.c. **analysis of the pros and cons of the technical characteristics** and **accuracy of the technologies to be compared (in relation to the main use at lab);**
5. Other instruments required during steps other than "sequencing" (see 4.b),

6. Options for purchase: plan and expected costs.
7. Costs of reagent & **average sequencing costs per genome (including personnel costs per genome sequenced!!!)**
8. Sharing the technology within the Institute: possible use by other departments and possible scenario and evolution for institutional purposes.

Critical points to be assessed

- **“Fragmentation and insertion of adapter sequences” into dsDNA**.
- **Depending on technologies it** may be **based on two separate steps (1. Fragmentation; 2. Insertion of adapters), or on a one-step process** (**tagmentation**), with enzymatic or physical methods.
- The one-step process has several advantages, including reduced sample handling and preparation time (personnel costs…).

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4351865/

- **Amplification of the tagged DNA**

Depending on technologies (and fragment size), you may need further instruments in order to select the DNA fragments of the desired size (e. g. you would need a Bioanalyzer for a routine use)

- **Sequencing**

After the library preparation and normalization steps:

Depending on technologies and instruments, before you perform your sequencing run:

-charging your analysis chip may be more or less user-friendly.

For accurate results you may need **additional instruments (up to +50% additional costs…)** and/or **more person/hours of skilled laboratory technicians per genome/per chip…**

Estimate total costs….

- Logistics (laboratory room, separate lab bench, temperature, humidity, internet connection, Uninterruptible Power Supply (UPS) … Data storage… Everything's under control?)
- NGS machine
- Other instruments needed beside the NGS machine
- Consumables: Reagents and kits
- Personnel costs

Example: estimated cost per genome (5Mb, coverage at least 50X, Nov 2016)

	Euro per genome (5 Mb)
Library preparation	32
Sequencing (best case X flowcell type)	41
Personnel cost (per genome)	9
Total (per genome)	82

Quantitative parameters to evaluate for the «sequencing output»

- **Accuracy**
- **Error rate**

Taking into account:

-Assessment of accuracy of the technology (e. g. literature on benchmarking and comparison with the «Sanger reference method»)

-Which microrganisms are mostly WGSed at you lab, and possible drawbacks when sequencing some types of microrganisms

Inspiring readings (at that time…)

Characterizing and measuring bias in sequence data" (Ross et al., 2013)
http://genomebiology.biomedcentral.com/articles/10.1186/gb-2013-14-5-r51

"Accuracy of Next Generation Sequencing platforms"(Fox et al., 2014)

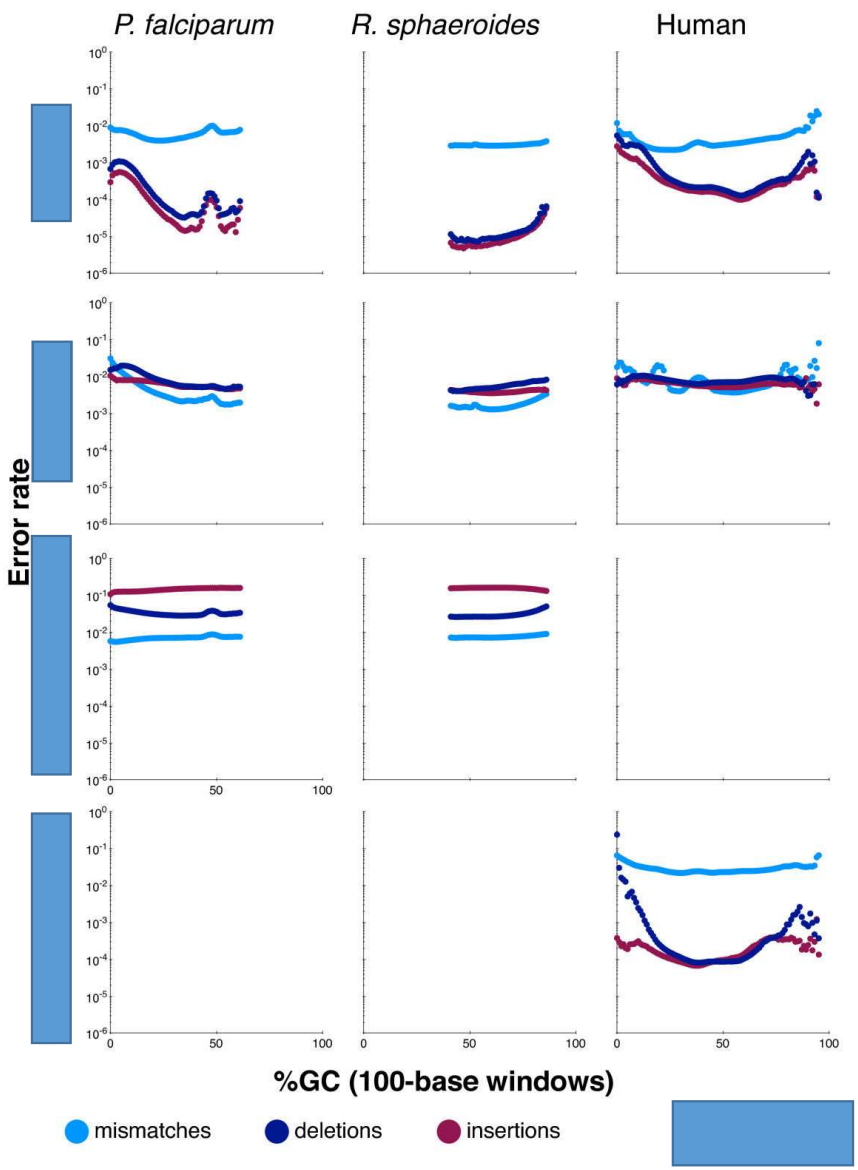
http://www.omicsonline.org/open-access/accuracy-of-next-generation-sequencing-platforms-jngsa.1000106.pdf

Characterizing and measuring bias in sequence data" (Ross et al., 2013)
http://genomebiology.biomedcentral.com/articles/10.1186/gb-2013-14-5-r51

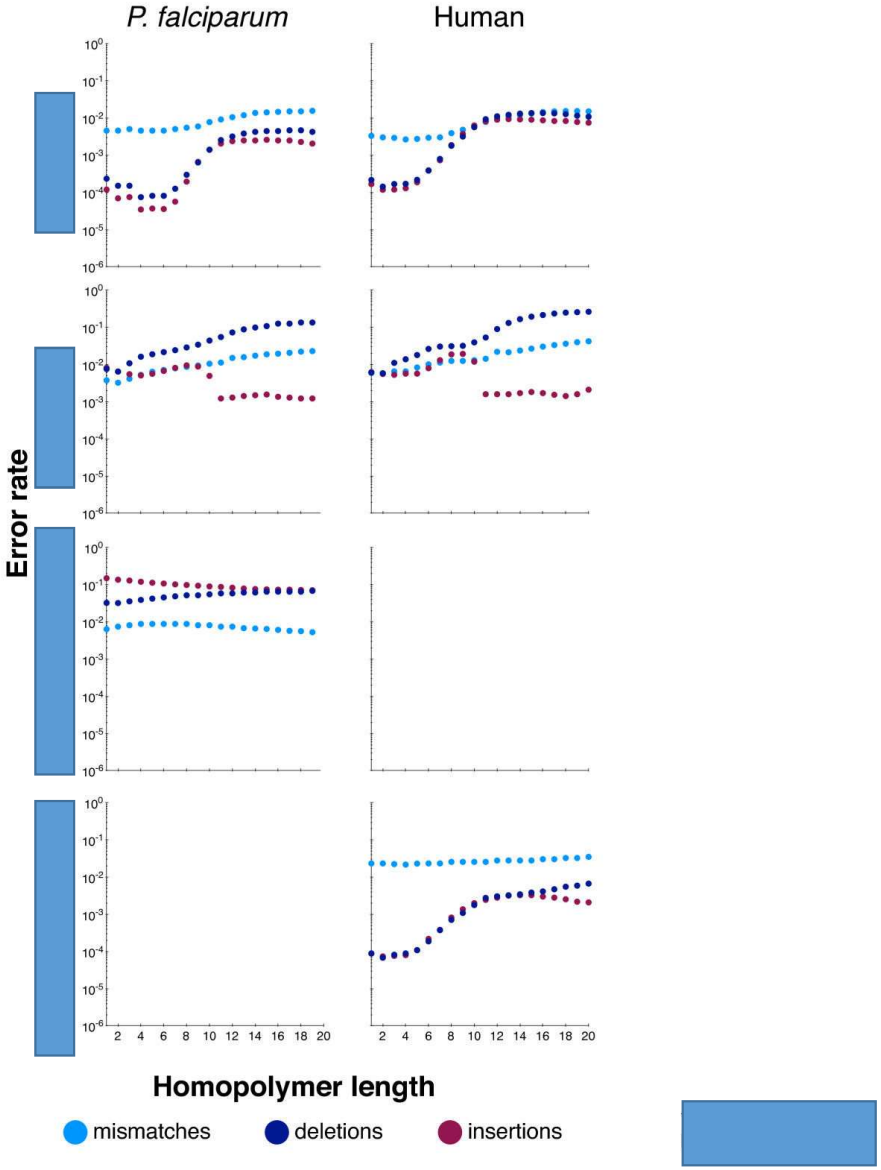
Sequencing technology error rates

Data set			Fractional error rate			
Sample	#	Platform	Mismatches	Deletions	Insertions	Total
P. falciparum	1	A(1)	0.0046	0.00021	0.00011	0.0049
	2	B	0.0038	0.0090	0.0068	0.020
	3	C	0.0068	0.033	0.14	0.18
E. coli	4	A(1)	0.0036	0.0000097	0.0000051	0.0037
	5	B	0.0018	0.0053	0.0044	0.012
	6	C	0.0077	0.032	0.17	0.21
R. sphaeroides	7	A(1)	0.0030	0.000018	0.0000089	0.0030
	8	B	0.0014	0.0055	0.0037	0.011
	9	C	0.0076	0.029	0.16	0.20
Human	14	A(2)	0.0030	0.00023	0.00017	0.0034
	15	B	0.0060	0.0069	0.0057	0.019
	16	D	0.023	0.000099	0.000091	0.024

Error rates as a function of GC composition



Error rates as a function of homopolymer length



In an article (Fox et al., 30 Apr 2014), it was **reported one or two log differences** in the error frequency (most frequent error type) among different technologies (10 ^-1 vs 10^-2 vs 10^-3)

http://www.omicsonline.org/open-access/accuracy-of-next-generation-sequencing-platforms-jngsa.1000106.pdf (2014)

Commercial Platform	Most Frequent Error Type	Error Frequency
A	single nucleotide substitutions	10-1
B	Deletions	10-2
C	CG deletions	10-2
D	Short deletions	10-2
E	A-T bias	2×10-2
F1	single nucleotide substitutions	10-3
F2	single nucleotide substitutions	10-3

All this background work generated a 25-page Summary Report…., and a purchase request

Istituto Zooprofilattico Sperimentale
del Lazio e della Toscana *M. Aleandri*

Direzione Operativa Diagnostica Generale

Centro di Referenza Nazionale per l'Antibioticoresistenza
National Reference Laboratory for Antimicrobial Resistance (Reg. 882/2004/EC)

Roma, 08/11/2016

Prot.

All.

Oggetto: Relazione per l'acquisto di apparecchiatura di "Next Generation Sequencing" per applicazioni di "sequenziamento di interi genomi" (incluso Whole Genome Sequencing).

La Struttura Complessa Direzione Operativa Diagnostica Generale, anche in qualità di Centro di Referenza Nazionale per l'Antibioticoresistenza e National Reference Laboratory for Antimicrobial Resistance (Reg. 2004/882/EC), e Centro di Riferimento Regionale per Agenti zoonosici Speciali, ha la necessità, per gli scopi istituzionali di acquisire uno strumento in grado di eseguire sequenziamenti massivi di genomi di agenti patogeni (batterici, micotici, protozoari) secondo tecnologia definita "Next Generation Sequencing" (NGS).

Finally, we made it!

Special thanks to all personnel at the Department, NRL-AR Italy