Current Trends in Microbiology
Vol. 10, 2016

Review

High-throughput DNA sequencing for genomic characterization of bacteria for monitoring hospital-acquired infections

Roberto Biassoni*, Eddi Di Marco, Alessandro Raso and Elisabetta Ugolotti
Istituto Giannina Gaslini, Genova, Italy.

ABSTRACT

Next-Generation Sequencing (NGS) technology allows to generate whole bacterial genome sequences in order to unequivocally characterize the strains to a highly discriminative resolution level. This methodology is used not only to study the bacterial clonality for tracing hospital outbreaks, but also to evaluate the presence of genes conferring antibiotic resistance or that encode for toxins. This makes NGS an appropriate technique that can be employed as 'standard practice' in hospitals, allowing to trace the circulation of "dangerous" strains. Here we describe the application of this technique in different clinical settings of suspected hospital-acquired infections sustained by relevant nosocomial pathogens. In particular, we analyze the more important genetic determinant and/or the mechanism conferring antibiotic resistance.

KEYWORDS: NGS, bacterial typing, WGS, nosocomial infections, outbreak

INTRODUCTION

Several typing techniques are in use that are useful in discriminating very closely related strains and such techniques are required in the control of hospital-acquired infections and for epidemiological purposes. For years such methods have been focused on band-based fingerprint techniques, such as pulse-field gel electrophoresis (PFGE) [1], amplified length polymorphisms (AFLP) [2], random amplification of polymorphic DNA (RAPD) [3], repetitive element Polymerase Chain Reaction (rep-PCR) [4], variable number tandem repeat (VNTR), and Multiple-Locus Variable number tandem repeat Analysis (MLVA) [5, 6] or Sanger sequence-based analysis like the single locus (SLST) [7] or multi locus (MLST) sequence typing [8, 9]. All these techniques are characterized by different resolution levels that sometimes give only limited information about the strains' relatedness. Hence, this has encouraged the development of novel methods, called Next-Generation Sequencing (NGS), that reach a greater genetic discrimination with low error rates and that give the potential to gather additional information. These platforms are able to produce high-throughput DNA sequencing of entire genomes at reasonable cost and time. Hence, the use of NGS is one of the most advanced approaches in the study of hospital-acquired infections. Thanks to an increase in the analytical resolution power, they are able to finely distinguish outbreak from non-outbreak isolates and to trace the route of strain transmission in clinical settings. This is the future trend since the bench-top NGS platforms make the whole microbial genome sequencing affordable and feasible and cheap enough for the clinical laboratory. At present, NGS applications are growing exponentially and

*Corresponding author: Roberto Biassoni,
Dipartimento Ricerca Traslazionale,
Medicina di Laboratorio, Diagnostica e Servizi,
U.O.C. Laboratorio Analisi,
IRCCS Istituto Giannina Gaslini,
Via Gerolamo Gaslini 5, 16147 Genova, Italy.
robertobiassoni@gaslini.org

there are "third-generation" platforms already available. These applications are able to analyze single DNA molecules producing very long DNA sequences that reach up to 10-20 kb, and some of the instruments used for such applications are so small that they can be used directly on field, without the need for fully equipped laboratories. At present, the field of whole genome sequencing is moving toward systems that are able to generate longer DNA sequences and characterized by lower error rates depending on the specific application. Thus, these recent technological advances make possible the study of whole bacterial genome useful to analyze local epidemiology, such as in the case of "intra-hospital outbreak studies", where it is necessary to know the microorganism genome at the single-base resolution. In addition, these analyses facilitate the rapid and accurate identification of pathogens, virulence and/or antibiotic resistance determinants, giving us the possibility to identify relationships between isolates and to group them into clonal lineages. Moreover, the analysis of similarities and differences among isolates allows us to understand the probable infection source, and the routes of pathogen transmission, and to develop the containment policy. Indeed, NGS technologies are essential tools that should be included in programs of microbiological surveillance in order to control the spread of infections and should be implemented by Hospital Infection Control Committees to understand how and why epidemic clones emerge, and thus to know how to manage them.

Classical typing methods

To date, several methods have been used for outbreak investigations and epidemiological surveillance of bacteria in clinical practice (summarized in Table 1). Many of those methods are based either on the genomic DNA analysis after restriction enzyme (RE) digestion or on PCR amplification analysis with or without RE digestion step. The most frequently used approach in outbreak screening is PFGE, which is based on the use of rare-cutter restriction endonuclease and 'pulsed-field' electrophoresis [1]. This "pre-PCR" method does not use any DNA amplification and/or automated sequencer analyzer with severe limitations in the portability of data. The introduction of PCR technique in the laboratory practice during the middle eighties gave rise to several applications based on amplification of sequence regions along the bacterial genome [10]. Initially, data detection was achieved by traditional gel electrophoresis but, over time, microfluidic capillary electrophoresis and/or Sanger-based sequencing were implemented and used, making both data collection and data portability easy. The PCR amplification of repetitive sequences followed by DNA electrophoresis is used in rep-PCR [4] and in MLVA typing [5] and in the subsequent evolution of these techniques [11] such as in DiversiLab system (BioMérieux Italia) and VNTR typing of *S. aureus* isolates [6]. All these methods are not "universal" since the amplification of targeted sequences for certain pathogenic species needs the design of dedicated primers. On the contrary, RAPD is an unspecific PCR amplification which uses short random oligoprimers and it does not need any previous knowledge about the genomic sequence of the pathogen. Although widely used for the typing of outbreaks, RAPD or its evolution like the "Arbitrarily Primed-PCR" [3] shows low reproducibility (inter- and intra-laboratory) and may lack the amplification of certain regions due to thermodynamical problems related to the low binding of the short random sequence primers as well as due to polymorphism in the priming binding site. PCR amplification followed by DNA digestion is used in AFLP [2]. The amplification is done using specific adaptors upon ligations and the detection of the amplified fragments is made possible using an automated DNA sequencer.

Multilocus sequence typing (MLST) is commonly based on the multiple locus allelic profile obtained by Sanger-based sequencing of 450-500 bp internal fragments of seven house-keeping genes. For each isolate the different house-keeping sequences define distinct alleles and the relative seven integer numbers of the seven loci characterize unambiguously the allelic profile or sequence type (ST). MLST and the relative databases have been created for the more common species of microorganisms of relevant clinical interest [8, 9] and could be easily accessed by Internet. It may be noted that this approach provides a portable, accurate, and discriminative typing system for macro-epidemiology analysis. Another highly accurate typing method (useful for certain pathogens) could be achieved by analyzing only the sequence of a specific and informative region of a single gene locus, for

Table 1. Traditional typing methods.

Typing system	Description	Purpose	Current use	Resolution level	Portability of data	Detection of unknown "sequence"	Plasmid detection	Limitations
Pulsed-field gel electrophoresis (PFGE)	DNA digestion and fragments separation by gel electrophoresis.	DNA fingerprint	Frequent	High, Poor (for bands with similar size)	Poor	Possible	Weak	Labour-intensive and time-consuming
Random amplification of polymorphic DNA (RAPD)	Unspecific PCR amplification. Separation by gel electrophoresis. (DNA sequence possible)	DNA fingerprint	Frequent	Medium	Poor Lack of inter-lab reproducibility	Possible	Possible	Low standardization
Repetitive-element PCR (rep-PCR)	PCR amplification of non-coding intergenic repetitive sequence and separation by gel electrophoresis.	DNA fingerprint	Infrequent	Pathogen dependent	Poor Lack of reproducibility	Low	Weak	Low standardization
DiversiLab system	Evolution of rep-PCR. Detection is obtained by high-resolution chip-based microfluidic capillary electrophoresis.	DNA fingerprint	Infrequent	Poor Limited bacterial species can be analyzed	Possible	Low	Weak	Limited to a restricted number of pathogen
Amplified fragment length polymorphism (AFLP)	DNA digestion and ligation followed by a specific PCR amplification. Detection is ensured by capillary electrophoresis separation.	DNA fingerprint	Frequent	Poor	Possible	Not possible	Possible	Labour-intensive and time-consuming
Single locus sequence typing (SLST)	PCR amplification and subsequent DNA sequence	DNA sequence	Frequent Limited to certain pathogen	Middle	Possible	Novel SL-type identification Possible	No	Limited to selected species
Multilocus sequence typing (MLST)	PCR amplification and subsequent DNA sequence	DNA sequence	Frequent	Middle	Possible	Novel alleles identification Possible	No	Limited to selected species

which freely accessible database is available. Indeed, it is known that accurate Methicillin-Resistant *Staphylococcus aureus* (MRSA) typing is possible using PCR amplification and DNA sequencing of repeat regions of the coagulase gene (coa) or the Staphylococcus protein-A gene (spa). The latter could be used for rapid typing of MRSA in the hospital setting and sequencing data could be analyzed using the Ridom StaphType web site (spaserver.ridom.de) [7]. The spa-typing is able to discriminate more than 16000 different sequences, thus giving a higher resolution than the coa-typing. This is possible since more than 700 different 24 bp-repeats may be assorted from a single to 27 repeats, thus generating the high degree of typing variability. In addition, spa-typing gives an optimal data portability like all the systems based on Sanger sequencing.

High-throughput methods

In the beginning, whole genome analysis of bacteria was achieved by the use of capillary electrophoresis equipment or automated Sanger-based sequencers. This sequencing approach gave in 1995, after one year of work, the first complete bacterial genome reported [12]. Later, in the same year, comparative genomic analysis of different bacterial genomes was also achieved for the first time [12, 13]. Today, the field of comparative bacterial genomics is moving a lot faster, since instruments able to produce many sequences in a few hours are available and a third generation of sequencers are already springing up.

Next-generation sequencers

Ten years after the first bacterial genome was published, Roche launched the first high-throughput DNA sequencing instrument (454), based on pyrosequencing technology [14]. Today, bench-top instruments such as Illumina sequencers (Illumina Inc, CA, USA), and Ion series (Ion torrent and Ion S5, Thermo Fisher, CA, USA) make the whole genome bacterial sequencing accessible and economically affordable for well equipped clinical laboratories [15]. These instruments utilize the massive parallel sequencing technology to generate up to millions of DNA sequences (reads) daily. The Illumina series instruments are based on the 'sequencing by synthesis' technology [16], while the Ion series use the semiconductor chip sequencing technology that is able to detect extremely small pH variations and does not require the use of optical signals.

The technology based on 'sequencing by synthesis' uses four fluorescently labeled deoxynucleotide (dNTP) terminators. First the templates immobilized on a proprietary flow cell surface are amplified (solid-phase bridge amplification) up to one thousand times in close proximity of the original DNA template to form a cluster of identical copies in a tiny area. This sequencing technology uses four fluorescently labeled nucleotides that work as reversible terminators in the polymerization reaction. At each sequencing cycle the correct dNTP is added, the fluorescence dye is detected to identify the base and later the fluorescent terminator is cleaved so that the sequence may continue with the addition of another single dNTP.

The Ion sequencers are based on a semiconductor technology which is able to detect the protons released during the sequencing reaction when nucleotides are incorporated [17]. In detail, DNA library fragments are linked by adaptors to the surface of beads particles and then amplified by emulsion PCR. The sequencing takes place in a chip containing millions of nano-wells where each well holds a single bead. The wells are on top of a semiconductor chip surface which is sensitive to proton flows. Then, the sequence is primed by specific adaptors that have been linked to the original DNA fragment during library preparation. The protocol is based on the sequential dispensation of the four bases. During the sequencing reaction the number of protons released is proportional to the number of specific bases incorporated and the pH change is detected in the well where the incorporation occurs.

The comparison between the two more common bench-top technologies has generated superimposable accuracy. Ion series are more prone to in-del errors (in homopolymeric regions), than Illumina, which is known on the other hand to produce higher frequency of base substitution errors [18]. Actually, these problems have been partly solved thanks to the advent of new high-fidelity chemicals and thanks to the improvement of the length of the sequences.

Third-generation single molecule real time sequencing (SMRT)

The SMRT instruments produced by PacBio and Nanopore are able to sequence the DNA at the single molecule level [15]. PacBio is based on a technology that uses fluorescence-labeled nucleotides [19], while the Nanopore detects an ionic current when a single-stranded DNA molecule goes through a 1.5 nanometer pore [20]. In detail the nucleotide incorporation momentarily blocks the DNA strand transfer through the pore and this causes a current flow change in a sequence/ nucleotide-dependent manner. Both PacBio and Nanopore technologies are able to generate very long sequences, for example the pocket-sized MinION platform (Nanopore) has been reported to produce more than 200 kb reads in length. Thus, these apparatus could be useful in studying complex haplotypes or in assembling whole genome sequencing with a small number of iterations [21, 22]. In the beginning, these novel technologies showed low accuracy, with an error rate that was about 14% for PacBio [23] and below 8% for Nanopore [24]. Improvements in their chemistries and in the purification protocols as well as of software routines have largely improved the accuracy of both SMRT systems such that the accuracy standards now have become more superimposable than the sequencers of second generation [25-27].

Bioinformatics

All bioinformatic approaches need to take into account that the bacterial genome is highly variable as it is continuously subjected to biological pressure due to immune responses and to antibiotic treatments. This variability is sustained by chromosomal mutation and also by the horizontal gene transfer. The analysis of whole bacterial genome requires a "*de novo* sequencing" approach based on short raw reads assembled by overlapping regions. To process Ion torrent sequencing data we tested four different softwares. Among them, Mira [28] and Spades [29] performed better than the Newbler [30] or Velvet [31], and moreover Spades is a lot faster than Mira to prepare assembled sequences. Although we find a high reproducibility in generating sequence contigs, it is known that some highly polymorphic areas, such as the Variable Number Tandem Repeat (VNTR) regions give problematic assembly. This problem is highly dependent on the length of raw reads assembled and it can be partially solved with longer reads. One critical example is the analysis of the *Staphylococcus aureus* protein A gene that contains a highly polymorphic region formed from a variable number of 24 bp-repeats. The typing (spa-typing) results from the combination of different polymorphic repeats, each of which is associated with a numerical code determined by an internationally available database (http://spa.ridom.de/spatypes.shtml). The numerically ordered combination of such repeats is used to define the spa-typing [7]. In our hands, the use of Spades assembler and the Hi-Q Ion Torrent chemistry gives > 60% precise typing results. It is to be noted that the information regarding the first sequence-repeat is often missing due to problematic assembly that causes the wrong attribution of the spa-type. Despite VNTR information in some cases are not attributed correctly, all together the thousands of the investigated genes used to define the strains' clonality are identified in a correct way, making the complete analysis reliable. This new approach allows the study of strains clonality in a better way than any classical typing methods. The study of phylogeny is carried out using the assembled sequences that are analyzed with a Single Nucleotide Polymorphism (SNP) variant caller in order to define the number of nucleotide differences for each strain compared to the one used as reference. These data could be used to generate a phylogenetic tree [32]. The easiest and fastest way to analyze the strain's parenthood is using the Ridom SeqSphere+ routine and database [33]. This software enables to automatically assess the classical MLST typing (based on the allelisms of 7 genes) and to analyze an extended number of bacterial genes covering the majority of genome sequences. In this way we easily reach an incomparable level of discrimination needed in the analysis of strain clonality. In addition, we may obtain useful information regarding all polymorphisms present in different genes.

Resistome and toxome

Next-generation sequencing technologies can also provide comprehensive information to define characteristic of microorganisms like antibiotic-resistance. Through this detailed analysis it is

possible to detect the presence of specific resistance genes, the presence of mutations responsible for the drug resistance and the regulatory mechanisms involved in gene expression (Resistome). Furthermore, you can also investigate if the resistance determinant is located on the bacterial chromosome or on mobile elements such as plasmids, and if resistance is vertically or horizontally transmitted [34]. In addition, by examining the trends and the prevalence of resistance determinants we may trace the spread of particular resistance factors in a bacterial population and their evolution. All these information are important not only for therapeutic purposes, but also for epidemiological aims. They allow to learn more about the local and global evolution of bacterial population, to trace clonal lineages and to follow international dissemination of antibiotic resistance [35]. The resistome analysis is newsworthy to be studied in multi-drug resistance (MDR) strains that are often involved in nosocomial infections and are difficult to be treated. Among these microorganisms a role of primary importance is certainly played by Extended-spectrum β-lactamases (ESBLs), carbapenemase-producing *Enterobacteriaceae* (CPE), Vancomycin-Resistant Enterococci (VRE) and Methicillin-Resistant *Staphylococcus aureus* (MRSA) [36, 37]. Today, these pathogens are the most challenging clinical problem because of their increased frequency of isolation in compromised patients and the difficulty in their treatment. In addition, the analysis of the whole bacterial genome allows to detect genes involved in metabolic pathways associated with the virulence and with the pathogenic factors (Toxome). For instance, our experience shows that in MRSA outbreak we could investigate different genes coding for different toxins such as the staphylococcal enterotoxins, the exfoliative toxins, the toxic shock syndrome toxin and the Panton-Valentine leukocidin (PVL) (lukS-PV and lukF-PV) [37, 38].

Practical application of whole genome sequence analysis by NGS

In the last few years we had focused on some clinically relevant microorganisms such as MRSA, *Acinetobacter baumannii* and *Enterobacter cloacae* for the investigation of suspected hospital-acquired infections. On this basis we studied these strains using Personal Genome Machine (PGM) like Ion Torrent Personal Genome Machine (Thermo Fisher Scientific, Carlsbad, CA, USA) and here we are reporting our analysis.

We tested and validated the use of NGS-based technology to assess Whole Genome Sequencing (WGS) in the management of nosocomial outbreaks, re-examining some informative MRSA strains that have sustained different clinical outbreaks. In detail, we studied 18 non-repetitive DNA isolates from infected patients of a pediatric intensive care unit (ICU) suspected to have sustained nosocomial infections. Thus, we have used PGM technology to define strain clonality, studying 1423 gene loci and we have analyzed the genetic determinants of antibiotic resistance (resistome) and genes encoding possible virulence factors such as toxins (toxome). Among the analyzed strains we found that 10 Neonatal Intensive Care Unit (NICU) strains were identical for more than 95% of the 1423 genes analyzed, and hence they could be defined clonally-related. In particular, by analyzing such strains in chronological order we understood that they had been responsible for three different rounds of hospital-acquired infections. Moreover, the analysis of two additional bacterial strains (sa43 and sa46) isolated from intensive care unit (ICU) indicated that they are identical for the allelism of 1422 genes out of a total of 1423 loci analyzed and thus were responsible for an additional "outbreak" (Figure 1, panel A). Therefore, NGS analysis allowed us to appreciate subtle genetic differences to classify closely-related strains in order to trace unambiguous parenthood of isolates and to define outbreak(s). In this way we have improved the genetic profiles obtained by Sanger sequencing (MLST, SpA and SSP-PCR for SCC-Mec) and PFGE typing. Finally, since we increased by a factor of 100 the number of gene loci analyzed, we now have the possibility to define the genetic information about the strains in depth and thus can establish the strain clonality more precisely. In addition, all phenotypic results of antibiotic susceptibility have been confirmed by the analysis performed using the WGS-assembled genotypes. Indeed, all MRSA strains clearly showed the presence of MecA gene; the fluoroquinolones (ciprofloxacin and levofloxacin) resistant strains had mutations of gyrA-S84L and grlA-S80Y; the aminoglycosides (gentamycin)

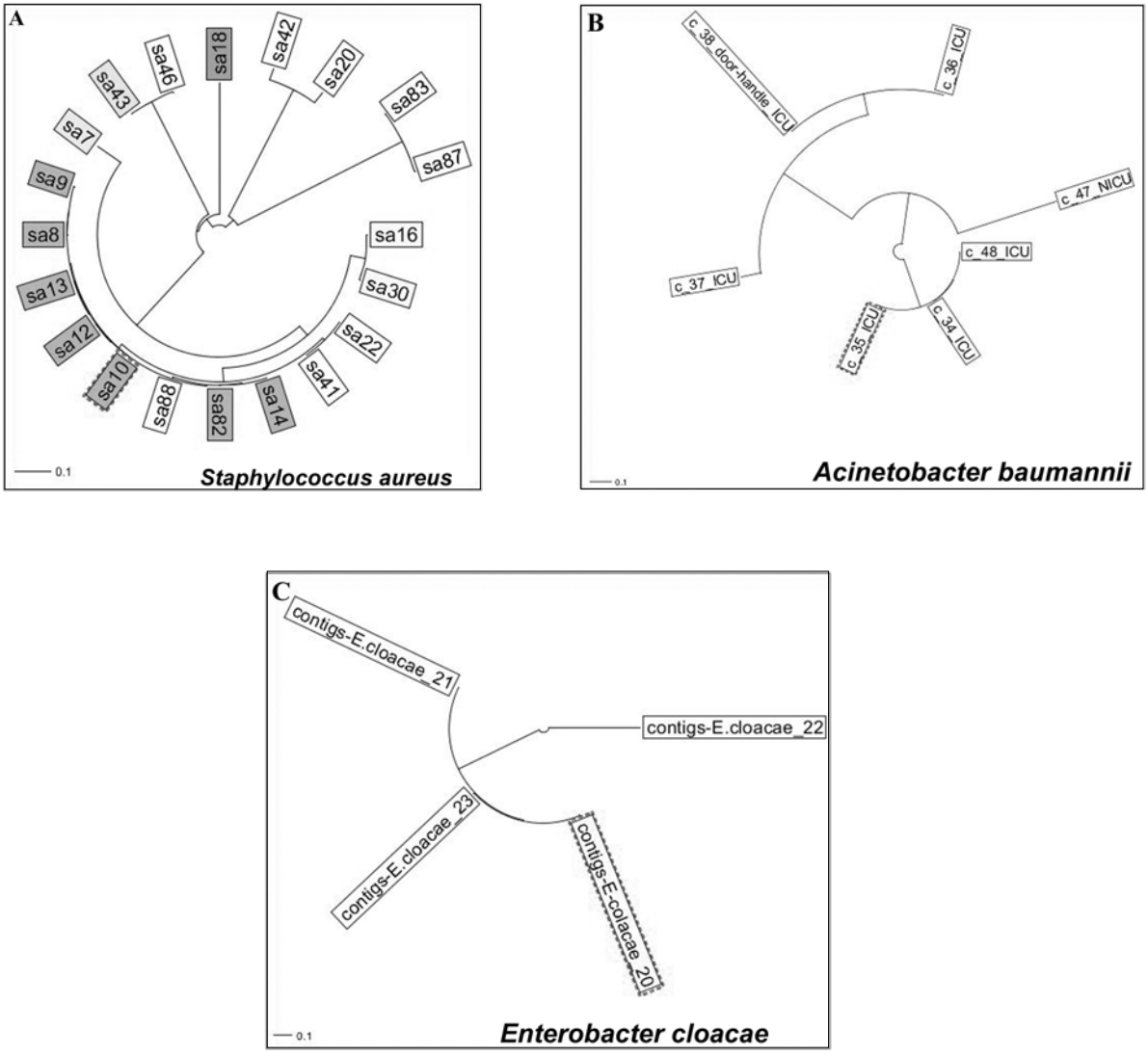

Figure 1. Phylogenetic tree of MRSA, *Acinetobacter baumannii* and *Enterobacter cloacae*. Figure 1 shows the circular Neighbor-Joining tree for **A:** MRSA, **B:** *Acinetobacter baumannii* and **C:** *Enterobacter cloacae*, to study the phylogenetic distance among different isolates.

resistant microorganisms presented the aacA-aphD gene. The resistance to Macrolides was sustained by the presence of ermC gene and trimethoprim resistance was related to the dfrA gene [36, 37, 39].

Among *Enterobacteria*, relevant to sustain hospital acquired infections, *Acinetobacter baumanni* plays an important role in view of its multi-drug resistance and its association to infections related to elevated morbidity and mortality. During the summer of 2015 we investigated a suspected outbreak that occurred in an ICU caused by *Acinetobacter baumanni-calcoaceticus* complex [40]. We analyzed 13 non-duplicated "*Acinetobacter calcoaceticus-baumannii* complex" strains isolated from 12 pediatric patients and from one environmental contamination discovered during a careful environmental monitoring promptly performed by the Infection Control Committee of our hospital. Also, in this case we sequenced the isolates using PGM platform that was demonstrated useful not only to determine the level of strain similarity, but also to identify the *Acinetobacter* species correctly in the first instance, since phenotypic and matrix-assisted laser desorption ionization (MALDI)

mass spectrometry data were inconclusive. Our analysis using the extended-MLST+ routine of Seqsphere software indicated that three A. baumannii strains are clonal, since they show more than 99.8% identical genes among the total number of 1701 genes studied. Moreover, we also identified an environment contamination since the two strains isolated, one from a patient and one from a room door handle, were 99.8% identical (Figure 1 panel B). These are important issues since using the information obtained by NGS we could enforce control measures in order to promptly limit any outbreak.

More importantly the use of NGS-based technology gave relevant information regarding resistome, since some of the analyzed strains that showed phenotypically multi-drug resistance to aminoglycosides, fluoroquinolones and β-lactams were genotypically confirmed. Indeed, we found genes coding for aminoglycosides-modifying enzymes like *aph(3")-Ib, aph(3')-Ic, aph(6)-Id*, *armA* 16S-RNA-methylase and mutations of *GyrA-S83L* and *ParC-S80K* genes that are known to determine a modified bacterial DNA-gyrase and topoisomerase IV enzymes, respectively. Moreover, *blaADC, blaOXA-66* and *blaOXA-23* β-lactamase genes, all under the control of the ISAba1 upstream genetic element, and the *blaTEM-19* are responsible for carbapenem resistance. Altogether these analyses confirmed the absence of genetic determinant for polymyxin resistance in accordance with antibiogram data.

Enterobacter cloacae represents the third type of nosocomial strains analyzed by NGS in our hospital. In detail, we collected four *Verona* integron-encoded metallo beta-lactamases (VIMs) positive *E. cloacae*. The infected patients belonged to the same department and all the bacterial strains were collected from rectal swabs within a limited period of time according to the Carbapenemase-Producing Enterobacteriaceae (CPE) surveillance. Also in this case, we have demonstrated that the strains were clonal (more than 99.4% identical) (Figure 1 panel C) and all showed *blaVIM-1* gene encoding for carbapenemases.

Final considerations

Advantages

The use of whole genome sequencing facilitates accurate identification of microorganisms thanks to an unequaled resolution of analysis compared to conventional typing strategies. WGS is able to yield objective and precise information, generating a large amount of data that are clinically relevant and easily interpretable by physicians. Additional benefit is the digital nature of sequencing data that allows sharing of information among national or international public health professionals. Moreover, WGS by the analysis of the complete pathogen genome gives a series of additional information about resistance and virulence mechanisms, relevant chromosomal gene mutation(s), presence of mobile genetic elements and medically important factors. Thus, this methodology results useful for surveillance/outbreak investigations and for the analysis of macro-epidemiology. In addition, NGS technologies enable to trace the geographical and historical evolution of multi-drug resistant pathogens responsible for infectious diseases and to alert healthcare staff about the diffusion of dangerous strains. Also WGS has the ability never shown before to define strain relationships, since it allows nucleotide-level analysis of the complete microorganism genome to discriminate outbreak strains from non-outbreak strains.

Limitation

In the beginning, the use of WGS technology was hampered because of the high cost of consumables and reagents. In addition, this technology requires personnel with both informatic and molecular biology skills. Today, thanks to novel instruments, new optimized chemicals and easier bioinformatic routines such limitations are becoming less relevant. In addition, NGS-based technologies are easily accessible to routine microbiology laboratories, thanks to a significant reduction in the cost of most of the reagents used. Moreover, the lack of bioinformatic expertise is partially overcome by the development of easily usable proprietary software suites and by the increasing number of web-accessible databases for the data analysis.

In conclusion, the significant impact of WGS technologies particularly in the hospital infection control may justify its implementation especially in diagnostic routine.

ACKNOWLEDGEMENTS

We would like to thank Gianni Berretta and Roberto Ceruti for taking care of the CLC-Bio Server and for Hardware support.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

REFERENCES

1. Tenover, F. C., Arbeit, R. D., Goering, R. V., Mickelsen, P. A., Murray, B. E., Persing, D. H. and Swaminathan, B. 1995, Journal of Clinical Microbiology, 33(9), 2233-2239.
2. Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J. and Kuiper, M. 1995, Nucleic Acids Research, 23(21), 4407-4414.
3. Barbier, N., Saulnier, P., Chachaty, E., Dumontier, S. and Andremont, A. 1996, Journal of Clinical Microbiology, 34(5), 1096-1099.
4. Versalovic, J., Schneider, M., de Bruijn, F. J. and Lupski, J. R. 1994, Methods Molecular Cell Biology, 5(1), 25-40.
5. Francois, P., Huyghe, A., Charbonnier, Y., Bento, M., Herzig, S., Topolski, I., Fleury, B., Lew, D., Vaudaux, P., Harbarth, S., van Leeuwen, W., van Belkum, A., Blanc, D. S., Pittet, D. and Schrenzel, J. 2005, Journal of Clinical Microbiology, 43(7), 3346-3355.
6. Sabat, A., Krzyszton-Russjan, J., Strzalka, W., Filipek, R., Kosowska, K., Hryniewicz, W., Travis, J. and Potempa, J. 2003, Journal of Clinical Microbiology, 41(4), 1801-1804.
7. Harmsen, D., Claus, H., Witte, W., Rothgänger, J., Claus, H., Turnwald, D. and Vogel, U. 2003, Journal of Clinical Microbiology, 41(12), 5442-5448.
8. Mayden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M. and Spratt, B. G. 1998, Proceedings of the National Academy of Science USA, 95(6), 3140-3145.
9. Spratt, B. G. 1999, Current Opinion in Microbiology, 2(3), 312-316.
10. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H. 1986, Cold Spring Harb. Symp. Quant. Biol., 51(Pt. 1), 263-273.
11. Healy, M., Huong, J., Bittner, T., Lising, M., Frye, S., Raza, S., Schrock, R., Manry, J., Renwick, A., Nieto, R., Woods, C., Versalovic, J. and Lupski, J. R. 2005, Journal of Clinical Microbiology, 43(1), 199-207.
12. Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, R. D., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J. F., Dougherty, B. A., Bott, K. F., Hu, P. C., Lucier, T. S., Peterson, S. N., Smith, H. O., Hutchison, C. A. 3rd and Venter, J. C. 1995, Science, 270(5235), 397-403.
13. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton G. G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O. and Venter, J. C. 1995, Science, 269(5223), 496-512.
14. Nyrén, P. 2007, Methods in Molecular Biology, 373, 1-14.
15. Pareek, C. S., Smoczynski, R. and Tretyn, A. 2011, Journal of Applied Genetics, 52(4), 413-435.
16. Reuter, S, Harrison, T. G., Köser, C. U., Ellington, M. J., Smith, G. P., Parkhill, J., Peacock, S. J., Bentley, S. D. and Török, M. E. 2013, BMJ Open, 3, e002175.
17. Rothberg, J. M., Hinz, W., Rearick, T. M., Schultz, J., Mileski, W., Davey, M., Leamon, J. H., Johnson, K., Milgrew, M. J., Edwards, M., Hoon, J., Simons, J. F., Marran, D., Myers, J. W., Davidson, J. F., Branting, A., Nobile, J. R., Puc, B. P., Light, D., Clark, T. A., Huber, M., Branciforte, J. T., Stoner, I. B., Cawley, S. E., Lyons, M., Fu, Y., Homer, N., Sedova, M., Miao, X., Reed, B., Sabina, J., Feierstein, E., Schorn, M., Alanjary, M., Dimalanta, E., Dressman, D., Kasinskas, R., Sokolsky, T., Fidanza, J. A., Namsaraev, E., McKernan, K. J., Williams, A., Roth, G. T. and Bustillo, J. 2011, Nature, 475(7356), 348-352.

18. Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L. and Law, M. 2012, Journal of Biomedicine and Biotechnology, 2012, 251364.
19. Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J, Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J. and Turner, S. 2009, Science, 323(5910), 133-138.
20. Stoddart, D., Heron, A. J., Mikhailova, E., Maglia, G. and Bayley, H. 2009, Proceedings of the National Academy of Science USA, 106(19), 7702-7707.
21. Koren, S. and Phillippy, A. M. 2015, Current Opinion in Microbiology, 23, 110-120.
22. Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R. Bertoni, A., Swerdlow, H. P. and Gu, Y. 2012, BMC Genomics, 13, 341.
23. Carneiro, M. O., Russ, C., Ross, M. G., Gabriel, S. B., Nusbaum, C. and De Pristo, M. A. 2012, BMC Genomics, 13, 375-382.
24. Jain, M., Fiddes, I. T., Miga, K. H., Olsen, H. E., Paten, B. and Akeson, M. 2015, Nature Methods, 12(4), 351-356.
25. Lin, H. H. and Liao, Y. C. 2015, PLoS One, 10(12), e0144305.
26. Goodwin, S., Gurtowski, J., Ethe-Sayers, S., Deshpande, P., Schatz, M. C. and McCombie, W. R. 2015, Genome Res., 25(11), 1750-1756.
27. Loman, N. J., Quick, J. and Simpson, J. T. 2015, Nat. Methods, 12(8), 733-735.
28. Chevreux, B., Wetter, T. and Suhai, S. 1999, Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB), 99, 45-56.
29. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A. and Pevzner, P. A. 2012, Journal of Computational Biology, 19(5), 455-477.
30. Kumar, S. and Blaxter, M. L. 2010, BMC Genomics, 11, 571.
31. Zerbino, D. R. and Birney, E. 2008, Genome Research, 18(5), 821-829.
32. Joensen, K. G., Scheutz, F., Lund, O., Hasman, H., Kaas, R. S., Nielsen, E. M. and Aarestrup, F. M. 2014, Journal of Clinical Microbiology, 52(5), 1501-1510.
33. Jünemann, S., Sedlazeck, F. J., Prior, K., Albersmeier, A., John, U., Kalinowski, J., Mellmann, A., Goesmann, A., von Haeseler, A., Stoye, J. and Harmsen, D. 2013, Nature Biotechnology, 31(4), 294-296.
34. Perry, J. A., Westman, E. L. and Wright, G. D. 2014, Current Opinion Microbiology, 21, 45-50.
35. Zarrilli, R., Pournaras, S., Giannouli, M. and Tsakris, A. 2013, International Journal of Antimicrobial Agents, 41(1), 11-19.
36. Ugolotti, E., Bandettini, R., Marchese, A., Gualco, L., Vanni, I., Borzi, L., Di Marco, E., Castagnola, E., Melioli, G. and Biassoni, R. 2011, Diagnostic Microbiology and Infectious Diseases, 69(2), 213-217.
37. Ugolotti, E., Larghero, P., Vanni, I., Bandettini, R., Tripodi, G., Melioli, G., Di Marco, E., Raso, A. and Biassoni, R. 2016, Journal of Hospital Infection, 93(4), 375-381.
38. Köser, C. U., Holden, M. T., Ellington, M. J., Cartwright, E. J., Brown, N. M., Ogilvy-Stuart, A. L., Hsu, L. Y., Chewapreecha, C., Croucher, N. J., Harris, S. R., Sanders, M., Enright, M. C., Dougan, G., Bentley, S. D., Parkhill, J., Fraser, L. J., Betley, J. R., Schulz-Trieglaff, O. B., Smith, G. P. and Peacock, S. J. 2012, New England Journal of Medicine, 366(24), 2267-2275.
39. Salipante, S. J., SenGupta, D. J., Cummings, L. A., Land, T. A., Hoogestraat, D. R. and Cookson, B. T. 2015, Journal of Clinical Microbiology, 53(4), 1072-1079.
40. Ugolotti, E., Di Marco, E., Bandettini, R., Tripodi, G. and Biassoni, R. 2016, Journal of Hospital Administration, in press.