Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Institut Pasteur
Research in Microbiology 165 (2014) 140–153
www.elsevier.com/locate/resmic

Meeting report

Microbial molecular markers and epidemiological surveillance in the era of high throughput sequencing: an update from the IMMEM-10 conference

The use of molecular markers to track bacterial, viral and other pathogenic agents as they spread and evolve is a key component of infectious disease surveillance and control. The International Meetings on Microbial Epidemiological Markers (IMMEM), initiated by the European Study Group on Epidemiological Markers (ESGEM) of the European Society for Clinical Microbiology and Infectious Diseases (ESCMID), are organized every three years to discuss progress in the field. IMMEM-10 was held on October 2–5, 2013 at the Institut Pasteur in Paris and gathered 400 participants from nearly 40 countries. The program included 190 posters and 72 oral communications, six of which were presented by industry. IMMEM-10 featured the most recent advances in the exploding field of pathogen genome sequencing and applications to public health for better prevention and control of communicable diseases.

1. Opening session

Sylvain Brisse (Institut Pasteur, Paris, France) provided historical background on IMMEMs and specific information about IMMEM-10, the first conference of this series to be organized in France (Fig. 1). Alex Friedrich, Chairman of the ESGEM group, reminded the audience of the goals and activities of this group within the ESCMID Society. Christian Brechot, President of Institut Pasteur, pointed out the close fit of the topics of IMMEM meetings with Institut Pasteur missions, including research in microbiology and genomics, as well as public health activities exemplified by the National Reference Centers hosted by the Institut Pasteur and the International Network of Pasteur Institutes.

Julian Parkhill (Wellcome Trust Sanger Institute, Hinxton, United Kingdom [UK]) illustrated how public health microbiology is in the midst of the revolution facilitated by high-throughput sequencing (HTS, also called next-generation sequencing [NGS]). He underlined current paradigm shifts in diagnostics, epidemiology and phenotypic knowledge of pathogens provided by real-time sequencing capacity. Taking the examples of *Staphylococcus aureus* and *Mycobacterium abscessus*, he showed how HTS provides resolution to discriminate between bacterial isolates even within short-term transmission chains. He also provided evidence that strain variation can accumulate silently in carriers or within patients, and discussed how this variation can confound the reconstruction of transmission events.

Andréa Ammon (European Centre for Disease Prevention and Control [ECDC] Stockholm, Sweden) discussed "The challenges of public health surveillance". She reminded us that the main 'raison d'être' of surveillance is to provide data that help define the best actions to prevent and control communicable diseases. She insisted that user requirements, not exciting technological possibilities, define the action priorities. A. Ammon presented the long-term surveillance strategic vision of the ECDC for the years 2014–2020, which aims to develop standards for surveillance systems, and discussed how technologies for pathogen characterization are just one of many aspects of surveillance, along with political, regulatory, ethical or legal aspects. Two tools for European-level surveillance were presented: TESSy (The European Surveillance System), a database of disease cases which now has a module to store information on molecular typing of isolates, and EWRS, the Early Warning Response System designed for early communication of public health events and related measures among member states. She also hinted at external sources of data that might be more integrated into surveillance in the future, such as Google Flu Trends.

IMMEM-10 was also an excellent occasion to honor two outstanding scientists, Mark Achtman and Brian Spratt, who have made multiple major contributions that have redefined the field of population genetics and molecular typing. Martin Maiden and Thierry Wirth retraced with panache these two exceptional careers.

2. Bioinformatics tools for genome-based microbial surveillance: practical demonstrations

Bioinformatics will play a central role in the era of HTS. Genome-based studies recently became much more affordable, providing enhanced discriminatory power for microbial surveillance. However, use of the bioinformatics tools needed to analyze these large amounts of sequence data is a challenge for surveillance laboratories and research laboratories alike.

0923-2508/$ - see front matter
http://dx.doi.org/10.1016/j.resmic.2014.01.003

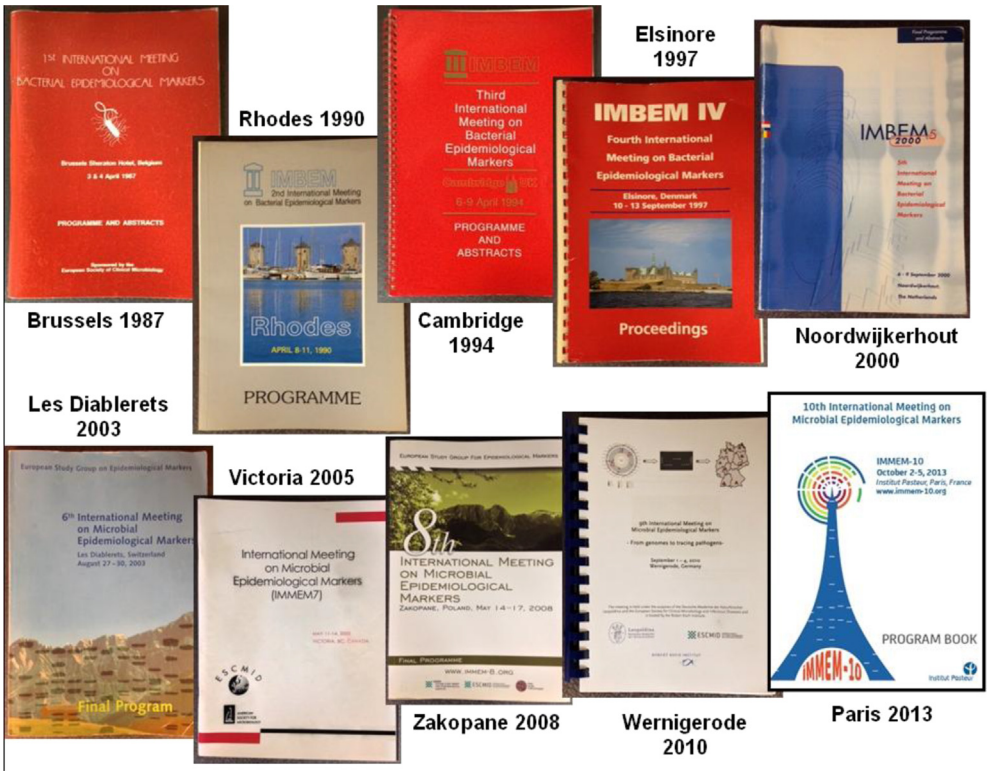

Fig. 1. The IMMEM-10 conference is the tenth edition of the successful series of meetings, called IMBEMs (for International Meeting on Bacterial Epidemiological Markers), until the year 2000. This figure features the cover pages of the program books of the ten meetings.

Dag Harmsen (Universitätsklinikum Münster, Münster, Germany) opened the session by giving an overview of concepts and available software tools for rapid HTS analysis in clinical and public health microbiology. He elaborated on the current switch from traditional typing tools, which are standardized and broadly used but provide only partial information, to more complete genome-wide approaches [23]. Genome-based approaches promise to increase the capacity to discriminate isolate genotypes to its theoretical maximum, providing possibilities that have been lacking thus far, e.g. for outbreak investigation. This change in scale not only refines the resolution of microbial characterization, but also opens up novel opportunities, such as identifying individual transmission events [24,36]. The challenges of HTS include the need for development of large storage capacity and databases and for standardization. Various genome-based characterization methods exist, from multi-species to lineage specific ones. The species-specific standard multilocus sequence typing (MLST) approach is expected to remain useful as strain nomenclature is well known and this provides a basis on which to build genome typing approaches [22].

In the next talk, Keith Jolley (University of Oxford, Oxford, UK) presented the Bacterial Isolate Genome Sequence database (BIGSdb). This database implements a gene-by-gene approach to genomic comparisons [25,41]. The system includes an isolate database where source information on isolates is stored. Any number of sequences, from single genes to complete genomes, can be linked to isolate data. The system also includes a sequence and profile definition database, which allows the grouping of loci into schemes. Schemes can include varying numbers of genes and range from MLST data to whole genome schemes. In addition, schemes with particular purposes such as virulence or resistance-associated elements or pan-genome schemes for a species can be defined, searched and tagged in a set of genomic sequences; this process was termed 'population annotation'. Additional BIGSdb functionality, called the Genome Comparator, enables rapid comparison of genomes based on user-selected schemes. Applications of this open-source software to molecular epidemiology were presented, demonstrating the benefits provided by the high resolution of HTS to population biology and outbreak investigation.

In his talk entitled "Phylogeographic reconstruction and visualization of pathogen spread through time and space", Philippe Lemey (Katholieke Universiteit, Leuven, Belgium) described a time-line for pathogen evolution. After reminding the audience that Ernst Haeckel used the phylogeography concept in the 19th century to explain the distribution of mankind on Earth, he described the added value of probabilistic model-based approaches for estimating the congruence of time and geography with phylogenetic reconstruction [5]. Several examples connecting either discrete or continuous traits to phylogenetics were presented, including the progression of West-Nile Virus in the USA, *Shigella* in Europe and the spread of influenza virus H1N1pdm [31]. The strong didactic power of the graphical representation of phylogeography, visualized in virtual globe software, was striking. Caveats of the approach and future perspectives were discussed and questions from the audience reflected the promise

and challenges of applying similar approaches to bacterial pathogens using complete genomes.

In the following talk, Eduardo Taboada (Public Health Agency of Canada, Lethbridge, Canada) discussed bridging the gap between whole genome sequencing and existing typing methods. He and his team have developed a tool for rapid generation of in silico typing data from draft microbial genome assemblies. MIST (molecular *in silico* typing) uses the BLASTN algorithm to integrate many diverse typing data. E. Taboada illustrated his presentation by an analysis of 2325 public and private *Campylobacter jejuni* draft genome sequences, and discussed concordance with existing typing methods [6].

Integrated platforms where genome annotation and comparison can be performed in parallel are of great help to guide annotation itself and are useful to link biology of the strains with their evolutionary success and epidemiology. This was the topic of the talk given by Claudine Médigue (CEA, Evry, France), who described "MicroScope", an integrated platform for the annotation and comparison of multiple microbial genomes. This platform first performs semi-automated annotation of bacterial genomes. The web interface, MaGe (Magnifying Genomes), allows user-friendly exploration and manual curation of the annotation results, providing the annotator with easy access to information derived from multiple resources, such as Uniprot, SwissProt, enzyme databases, and synteny views with user-defined reference genomes. In recent years, the platform has evolved to integrate other functionalities including genomic island detection, sequence variation and small-scale evolution, transcriptome analysis based on RNAseq data, and metabolic reconstruction and network comparison [45].

Gegenees, a software tool for comparing hundreds of genomes, was presented by Anders Sundström (National Veterinary Institute [SVA], Uppsala, Sweden). At the core of this approach, the genome is split into chunks after which the BLAST algorithm is used to compare genomes. Gegenees allows identification of unique genome sequences specific for subsets of strains as well as the visualization of common strain-specific genomic regions [2].

One important aspect of genome comparisons for epidemiological purposes, especially at the fine temporal scale of outbreaks or for highly monomorphic pathogens, is to maximize the amount of information included in the analysis while optimizing its reliability. Hannes Pouseele (Applied-Maths, Sint-Martens-Latem, Belgium) presented his work entitled "Optimized whole genome MLST [wgMLST] versus whole genome SNP [wgSNP] typing in NGS-based epidemiological surveillance". As wgMLST uses defined loci, which generally only include the coding sequences, intergenic regions are not considered; therefore, important single nucleotide polymorphisms (SNPs) may be overlooked. Comparing 26 *Mycobacterium tuberculosis* genomes, the authors found 347 SNPs by a whole-genome mapping approach, with some regions of the genome being unmappable due to repeated sequences. wgMLST included 4018 genes, but 47 were removed as they included repeated sequences, which led to difficulties in assembly, mapping and/or allele calling. Further, different calls of alleles were found in some pairs of strains believed to be identical, which was attributed to paralogous sequences. Although this issue could be solved in the future using genomic localization information, nearly 10% of the genes were removed from the analysis. Overall, the resolution of wgMLST was found to be similar to wgSNP.

Continuing the exploration of the potential of HTS microbiological characterization, João Carriço (Instituto de Medicina Molecular, University of Lisbon, Lisbon, Portugal) reported on an HTS pipeline process using an ontology approach and its application to molecular epidemiology. NGSOnto proposes to summarize the historical process through which the data were produced, aiming to keep track of the whole process and to provide a universal resource for enabling comparison among independent datasets and studies. NGSOnto was designed using Web Ontology Language (OWL) and is based on the existing data formats of NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) and EMBL-EBI European Nucleotide Archive (http://www.ebi.ac.uk/ena/). All steps, from DNA extraction to the final sequence, are recorded in the NGSOnto database and can be retrieved by using SPARQL queries (http://www.phyloviz.net/NGSonto/). Therefore, NGSOnto will provide the community with a machine-readable format for querying of multiple websites, comparing datasets and creating interoperable databases for microbial typing and phylogenetic analyses.

3. Outbreak genomics and epidemiology

HTS is transforming our ability to track outbreaks and make epidemiological inferences. The Outbreak Genomics and Epidemiology session covered a wide range of topics along this theme, from the practicalities of integrating whole genome sequencing into existing infrastructure, to an example where inference from genome sequences conflicted with national health policies. Jonathan Green (Public Health England [PHE], Colindale, UK) described progress in establishing a centralized sequencing service at PHE to monitor and track outbreaks of infectious diseases. Although HTS technology is rapidly evolving, PHE believes that current technology is robust enough to make investment worthwhile, even though costs are very large and an overhaul of the existing infrastructure is required. The need for robotics, a specialized staff, bioinformatics support, expensive hardware as well as training services was underlined. Priority areas were identified, such as emergency response and a small number of pathogens of high public health importance. Using *Streptococcus pneumoniae* as a case study, J. Green asked whether HTS could replace the current serotyping system. Whilst there is a very high correlation between sequence and serotype, a small number of cases are hard to resolve, but this should hopefully change in the future by building a bank of reference serotypes. This early experience was judged very positive, but automation and validation of this novel approach are likely to take time and to bring up many new issues such as data sharing and the ethics of using HTS data in legal cases.

UK government sanctions have been put into place to fine hospitals with large numbers of *Clostridium difficile* cases,

using the reasoning that poor hospital hygiene is the causal problem. Using 957 genomes from Oxfordshire infection cases over a 3.6-year period, Tim Peto (University of Oxford, Oxford, UK) and colleagues examined the relationships among isolates in individual hospitals [9]. A genetic threshold of epidemiological linkage was defined based on the whole dataset. Using this threshold (>10 SNPs), the authors showed that 45% of cases were not epidemiologically related and that only 35% could be conclusively linked. Of the latter, 38% could be explained by patterns of patient contact. These results imply that the majority of cases may be unrelated and are actually coming from asymptomatic carriage or some unknown environmental reservoir. *C. difficile* infection may therefore be a wider public health problem, with within-hospital transmission playing a smaller role than current consensus opinion suggests.

Remaining on the topic of *C. difficile* infection, Mathias Steglich (Robert Koch Institute, Wernigerode, Germany) described the use of HTS in analysis of a particular ribotype (027/NAP1/FQR) in Germany. There has been an increase in *C. difficile* infections in Germany over the last decade, since an aggressive outbreak in 2007 [40]. The most recent common ancestor of these strains was dated back to 2001 for fluoroquinolone-resistant variant 1 (FQR1) and 2003 for FQR2, implying that these strains had been circulating for some time before the first recognition of an outbreak. The spatial dynamics of the HTS data concurred with epidemiological information, demonstrating the power of this approach in tracking disease transmission.

Eija Trees (Center for Disease Control [CDC], Atlanta, USA) investigated Shiga toxin-producing *Escherichia coli* (STEC) isolates from sporadic cases and several outbreaks in the USA. In this study, 237 STEC genomes were sequenced in order to assess congruence between HTS data, epidemiological data and pulsed-field gel electrophoresis (PFGE), which is currently the standard typing method. The outbreaks themselves were temporally diverse; some occurring over just a matter of days, with others lasting many months, which was reflected in the genetic diversity. Whereas total genetic variation was large (>10,000 SNPs), often only a few SNPs distinguished the different isolates within an outbreak. However, in one child patient alone, a diversity of 47 SNPs was detected over a 2.5-month period. Overall, HTS showed a good correlation with PFGE, but resulted in a much more detailed picture of the outbreaks.

Christoph Eller (Robert Koch Institute, Germany) described a rather distressing outbreak of nosocomial *Klebsiella pneumoniae* infection in a neonatal intensive care unit in Bremen, Germany. The ward remained closed following 31 cases and five deaths in 2011–2012 as the infection persisted even after closure and renovation of the ward in January 2012. Using HTS of 37 isolates, C. Eller attempted to track the source of the outbreak. The most recent common ancestor of the outbreak was estimated to have existed in 2008, implying that strains were circulating a few years before the outbreak. Patient to patient transmission was the likely cause of maintenance of the infection, and the persistence of infection during the renovation remains a mystery. This study highlighted the need for greater sampling during an outbreak and a better understanding of the natural history of microbes, i.e. survival in the environment and asymptomatic carriage. It was concluded that without this knowledge, HTS could not reach its full potential.

4. Population genetics and pathogen emergence

This session addressed the emergence of human pathogens, including multidrug-resistant lineages, at different temporal scales - from historical times to contemporary epidemiological events. In his talk, Martin Maiden (University of Oxford, UK) promoted a 'gene-by-gene' approach using either 53 ribosomal genes or core genome MLST to reliably resolve *Neisseria* species and their distinct lineages, such as hyperinvasive meningococcal clones or gonococcal lineages [3,34]. Core genes, accessory genes and those that may be associated with particular lineages or increased pathogenicity potential could also be identified, for example in the globally distributed ST-32 clone. The MRF Genome library (http://www.meningitis.org/current-projects/genome), an example of genome database that will help future strain comparisons, was also presented.

Stefan Niemann (National Reference Centre for Mycobacteria, Borstel, Germany) reported on genomic epidemiology studies of tuberculosis (TB). Understanding transmission is crucial for control of TB, a major public health issue that accounts for nine million cases per year worldwide. Among these, approximately 0.5 million cases are caused by multidrug-resistant TB (MDR-TB). HTS is particularly suited to longitudinal epidemiological study of TB due to the high genetic homogeneity of *M. tuberculosis*. Estimated times and patterns of divergence of *M. tuberculosis* lineages mirror those of human mitochondrial DNA (mtDNA) and the 'Out-of-Africa' model of evolution of modern human populations [7]. HTS can also be applied to estimate the efficacy of therapy and to track global spread of strains and transmission dynamics. S. Niemann concluded his talk by stressing the urgency of integrated standardized tools, including databases and nomenclature, if NGS is to become a standard in clinical microbiology laboratories.

TB was also the subject of the next talk, given by Thierry Wirth (Muséum National d'Histoire Naturelle, Paris, France). MDR and extended multidrug-resistant (XDR) lineages, including the Beijing lineage, are increasing in prevalence in France [4] and other parts of Europe. Most cases of MDR-TB come from eastern Asia, where the largest diversity is found. Using coalescent-based analyses, T. Wirth showed that changes in the effective population size of *M. tuberculosis* over time mirror changes in human demography and were impacted by major events such as the Industrial Revolution, World War I, use of antibiotics and the emergence of HIV. He also showed how the Beijing lineage has evolved over time from its low-virulence ancestral form through the industrial age and antibiotics era to become increasingly virulent and resistant.

5. Molecular typing and epidemiology

The development, evaluation and application of molecular markers have been the central topic of IMMEM conferences since the first IMMEM in 1987. In his talk entitled 'Forget your gels: Population genomics versus fingerprinting', Mark Achtman (Warwick University, Warwick, UK) described the current state of molecular typing in bacteria and offered advice about the best techniques to use, with a focus on the reliability of strain relationship inferences, population genetics studies and long-term epidemiology. As one of the inventors of MLST, M. Achtman reminded the audience about his early warning against problems associated with developing yet another typing method (YATM) and using 'tried and true but stodgy methods' (TATBSTMs) such as ribotyping, serotyping, AFLP, RAPD, PFGE and MLVA. Based on the eloquent example of *Salmonella*, it was shown how most of the >1500 serovars are probably polyphyletic (i.e. have multiple founding ancestors), due to horizontal genetic exchange and environmental selection for certain serotype antigens. In spite of this, serotyping remains the most commonly used *Salmonella* typing method in public health laboratories. M. Achtman strongly advocated forgetting serotyping of *Salmonella*, given its poor phylogenetic value and other limitations [1]. The power of genomics to investigate outbreaks and genome dynamics within a single serovar was then illustrated on *Salmonella enterica* serovar Agona, which caused multiple outbreaks in recent years. In particular, it was shown that identical types that are defined based on SNPs can have very different PFGE patterns because of bacteriophage integration [48], questioning the reliability of PFGE as a strain typing method.

Rob Willems (Utrecht University, Utrecht, The Netherlands) reported on the high-resolution molecular epidemiology of vancomycin-resistant *Enterococcus faecium* (VRE) during an epidemic in Dutch hospitals. Typing 257 VRE isolates from 27 hospitals in the Netherlands by MLST revealed 26 sequence types (ST), among which 12 were found in more than one hospital. In nine hospitals, more than two different STs were found. This indicates that the epidemic rise was polyclonal and may imply spread among hospitals. STs were distributed among three clades, which could be broadly classified as containing (i) human commensal strains; (ii) animal-associated strains; and (iii) nosocomial strains. Whole genome sequencing of 25 isolates with varying vancomycin resistance levels was consistent with an evolutionary scenario whereby lineages originating in animals, not from human commensal carriage, entered the hospital environment and are circulating. This was associated with the acquisition of different *van* genes through horizontal gene transfer and followed by the epidemic rise of VRE in the Netherlands.

Joakim Ågren (SVA, Uppsala, Sweden) reported on the investigation, based on six whole genome sequences, of a Swedish cattle disease outbreak of *Bacillus anthracis*. Given that the outbreak was short-lived, confined to a specific area and had a single source, and because *B. anthracis* is a genetically monophyletic species, whole genome sequencing was particularly useful to detect SNPs and short insertions and deletions (indels). Isolates from different animals typically differed in ~1–3 SNP or indels. Antibiotic resistance tests showed that isolates from animals that had died during or after penicillin treatment of the cattle had developed penicillin resistance. The resistant isolates had independently mutated the same regulatory genes, thereby activating a resistance gene. The study also revealed heterogeneity of clones within the same animal. Because this heterogeneity was as large as that between different animals, the use in epidemiological investigations was limited and it was difficult to pinpoint the exact source of the outbreak.

Focusing on *Staphylococcus capitis*, one of the most common coagulase-negative staphylococci responsible for infant death in neonatal intensive care units (NICU), Patricia Martins Simões (Centre National de Référence des Staphylocoques, Lyon, France) presented a study based on PFGE, MLST, dru-typing and SCCmec typing to investigate population structure and multidrug resistance in an international set of NICU *S. capitis* isolates from Australia, Belgium, France and United Kingdom. There was little genetic variation among *S. capitis* from neonatal samples when compared to isolates from adult bacteremia, and all neonatal samples had the same PFGE profile and nearly identical MLST profiles, dru- and SCCmec-type, thus seemingly belonging to a single clonal population with world-wide distribution (clone NRCS-A). The SCCmec element of a prototype strain from this clone was fully sequenced and compared with four other whole-genome sequences of *S. capitis*. The authors put forward an evolutionary hypothesis involving two independent SCC element acquisitions (an SCC element harboring genes of resistance to heavy metals and an SCCmec element) and clonal expansion in the French NICU environment.

6. Phylodynamics of viral pathogens

Combining phylogenetics, population dynamics and geographical dispersion analyses results in a powerful framework for reconstructing the temporal-spatial dynamics of viral epidemics. In a talk entitled "Phylogeodynamics of HIV", Anne-Mieke Vandamme (Katholieke Universiteit Leuven, Belgium) reviewed knowledge on the origin and spread of HIV1 and HIV2. Their initial evolution from simian immunodeficiency virus (SIV), following cross-species transmission was estimated at the end of the 19th century and the beginning of the 20th century for the major groups. The study of strain diversity in humans revealed how several subtypes of HIV subsequently arose from founder events and how social factors, sexual practices, geographical constraints and political forces influenced their diversification and spread. As an example, dispersion of the HIV1-C subtype was linked to the copper belt corridor (southern and southeastern Africa), whereas HIV1-B migrated from the Congo to the US through Haiti in two main historical steps: after the Congo acquired independence in the 1960 and during the gay movement emancipation in the 1970's [14]. Likewise, one HIV2 lineage went through a founder effect out of the Côte d'Ivoire, with a probable effect on the HIV 2 epidemic of socio-cultural changes induced by the war

of independence in Guinea-Bissau [30]. During the discussion, it was explained how phylogenetic analyses led to identification of putative adaptations permitting human infection, while experimental work using cell culture provided further insight into the molecular mechanisms at work.

The globalization of food markets and of international travel promotes the spread of pathogens. Rita de Sousa (Dutch National Institute for Public Health and Environment [RIVM], Bilthoven, The Netherlands), discussed how geotagging of hepatitis A virus (HAV) represents a dynamic tool for public health surveillance. HAV is a food-borne virus endemic in several world regions. Three main genotypes are currently described as causing human infection. The objective of this study was to analyze the suitability of the international hepatitis A database (HAVNET) for detection of international foodborne outbreaks. The database contains data from over 7000 HAV sequences detected between 1957 and 2013, of which 75% could be assigned to the most probable country of origin of HAV. Using the available data in HAVNET, it was possible to identify and link outbreaks with their geographic origin, although comparative analyses of sequence data is hampered by lack of worldwide genotyping standardization. R. de Sousa discussed the intrinsic limitations of such international collaborations regarding access to virological or epidemiological data. In addition, the need for genotyping standardization and for increasing the geographic representation and coverage of data was emphasized in order to improve the applicability of the database to support public health investigations.

The last talk of this session, entitled 'Phylogenetic patterns of human coxsackievirus B5 arise from population dynamics between two genogroups and reveal evolutionary factors of molecular adaptation and transmission' was given by Jean-Luc Bailly (Université d'Auvergne, Centre National de Référence des Entérovirus/Parechovirus, Clermont-Ferrand, France). Coxsackievirus B5 (CVB5), a serotype of the enterovirus B species, is mainly involved in mild infections that can, in some cases, induce fever, meningitis, and even myocarditis. The occurrence of this common enterovirus follows a seasonal epidemic pattern. To gain insights into the tempo and mode of evolution that sustain the genetic diversity of CVB5 and its interplay with virus transmission, 218 virus samples were investigated. Phylogenetic reconstructions showed two distinct CVB5 genogroups. Genogroup B was estimated to have been co-circulating with genogroup A over the last 50 years. A coalescence study revealed three peaks of genetic diversity, in 2000, 2003 and 2006. Ancestral state reconstruction revealed that the evolutionary processes responsible for the diversity of viral protein 1 were different between the two genogroups: in genogroup A changes were distributed all along the phylogeny, while in genogroup B all changes were mapped at the origin of the clade. These differences were attributed to the immune selective pressure of the host. Finally, it was suggested that environmental samples and pediatric as well as asymptomatic cases should be considered for a more complete understanding of the epidemiological pattern of enterovirus infections.

7. Practical use of typing methods in epidemiology

Typing methods should have three important properties that favor their wide application: high discriminatory power, simplicity of use and low cost. In light of the whooping cough resurgence in several countries, Norman Fry (PHE, Colindale, UK) reviewed existing recommendations for *Bordetella pertussis* strain molecular characterization [37]. The observed increase in infection could reflect waning immunity, improved surveillance or pathogen adaptation. Understanding these factors is an important public health objective. Of the typing methods that are used, serotyping offers low discrimination, but provides an easy way of detecting changes in *B. pertussis* populations. PFGE, MLVA and DNA sequencing offer higher discrimination. More importantly, DNA sequence-based methods can monitor changes in particular genes – for example genes coding for cellular vaccine components, but are more labor-intensive. As for other microorganisms, whole-genome sequencing offers the greatest amount of information. For monitoring of the emergence of vaccine antigen-deficient variants, N. Fry and colleagues recommend a combination of DNA sequence analysis (Sanger or HTS) and protein detection techniques, which must ideally conform to guidelines for the validation and application of typing methods used in bacterial epidemiology [46].

In the next talk entitled "*S. aureus* USA300: how to discriminate between the bad, the worse and the ugly", Corinna Glasner (University Medical Center Groningen, Groningen, The Netherlands) presented the practical use of a genome-based typing method for the easy, rapid and low cost identification of different *S. aureus* types. C. Glasner described a study of nosocomial and community-associated MRSA isolates, all featuring the USA300 PFGE profile, but two different and highly-related *spa* types, using multiple-locus variable number of tandem repeat fingerprinting (MLVF). In addition, the patterns of exoproteomes of the typed isolates were compared in a search for novel microbial epidemiological markers. MLVF grouped the isolates into three distinct clusters. Importantly, it was clearly shown that MLVF was capable of overcoming the limitations of PFGE in distinguishing highly related *S. aureus* clones [15,28]. Differences in the secreted proteins matched the three MLVF clusters, reflecting epidemiological relationships among the respective isolates. It was concluded that MLVF can be useful for epidemiological investigations, outbreak prevention and control, especially in countries where HTS is not yet in reach.

Rebecca Lindsey and John Besser (CDC, Atlanta, USA) presented 'New Tools to Investigate Outbreaks: Shoe Leather Molecular Epidemiology'. Molecular typing methods such as PFGE and MLVA are still widely used at the CDC and are important for understanding foodborne disease epidemiology. The widespread use of molecular tools has raised scientific and legal issues and led to thinking about disease transmission in new ways. Novel sequencing methods will need to be evaluated and adapted in order to fully exploit their potential for molecular epidemiology purposes. This

will require statistical evaluation of cluster significance, fluid case definitions for maximizing the signal-to-noise ratio and analyses that integrate information on the microbial agent with person-place-time dimensions. The talk concluded with a discussion on the continued need for interpretation by skilled epidemiologists, with a focus on food-borne disease epidemiology.

8. Virulence emergence, diagnosis and epidemiology

Defining the molecular factors that lead to virulence is an important research target. HTS offers novel opportunities, since comparisons among genome sequences of isolates causing different clinical outcomes or pre- and post-emergence can help in identifying virulence-associated genetic factors. The role of phages and other horizontally transferred genetic elements in the adaptation of bacteria to novel environments was a key theme in this session. However, it was also pointed out that identifying virulence factors from whole genome sequences can be complex, and quantifying the different effects of the bacterium and host in the manifestation of virulence is needed. Siv Andersson (Uppsala University, Uppsala, Sweden) provided a genomic perspective on the adaptive radiation of the emerging pathogen *Bartonella* [20] in her talk "The ecological context of pathogens and the evolutionary dynamics of their virulence genes". There are three distinct phylogenetic clades of *Bartonella*, each roughly associated with a different host type: a ruminant group; a cat and rodent group; and a human, cat, dog and rodent group. Adaptation to humans appears to have evolved twice. Mapping the dynamic gene flux of *Bartonella* across the species phylogeny, Andersson's group identified a gene transfer agent (GTA) containing a phage-derived origin of replication, located near a secretion system cassette. This GTA is being transferred between different strains of *Bartonella*, suggesting a role in host adaptation. An example of transfer of the secretion system was found between a dog-adapted strain and a cat-adapted strain. The dissemination of host-specific genes could have been important in the successful adaptation of *Bartonella* to different host groups. In humans, modification of the existing gene pool through gene loss was observed, but no novel human-specific gene was found, which may be explained by the high genetic homogeneity of the population.

In the second talk of the session, Ulrich Dobrindt (Universitätsklinikum Münster, Münster, Germany) outlined the complicated pathogenic dynamics of *E. coli* isolates. Uropathogenic *E. coli* (UPEC) are often phenotypically and genotypically distinct from commensal *E. coli* isolates. However, some are superficially indistinguishable from commensals, and others even resemble intra-intestinal pathogenic *E. coli* (IPEC) or other extra-intestinal pathogenic *E. coli* (ExPEC). This highlights the need for better typing of *E. coli* isolates. U. Dobrindt showed that MLVA offers better resolution than MLST, but it is still unable to ambiguously distinguish the different pathotypes. The distribution of genes encoding autotransporter proteins, known to contribute to virulence, was determined from the sequencing of 111 whole genomes. Whilst their distribution correlates with phylogeny, these genes also do not appear to be reliable biomarkers of pathotypes. In fact, there was little pathotype distinction even using whole genome sequences, perhaps highlighting the specific host contribution to clinical outcome.

Carmen Buchrieser (Institut Pasteur, Paris, France) reported on the population genomics and evolution of virulence in *Legionella*. This genus is responsible for outbreaks of acute respiratory illness in humans, as it is transmitted by inhalation of aerosolized contaminated water. Natural reservoirs of *Legionella* are amoebas from lakes and rivers. Serogroup 1 of *Legionella pneumophila* alone is responsible for 90% of infections. C. Buchrieser showed that the different serogroups did not cluster together in a phylogeny based on whole genome sequences. This is attributed to horizontal transfer of the lipopolysaccharide cluster between genetically different strains. In France, two serogroup 1 clones are responsible for 30% of disease cases. Interestingly, these clones have horizontally acquired a thiocyanate hydrolase gene. As thiocyanate compounds are used to clean water circuits, this acquisition may represent an adaptation to this specific niche [16].

Keeping on the topic of virulence emergence, Lucy Weinert (University of Cambridge, Cambridge, UK) reported on the genomic differences between commensal and pathogenic *Streptococcus suis*. *S. suis* causes respiratory and systemic disease in pigs, but is also zoonotic and the leading cause of human meningitis in Vietnam. Using genome sequences from 400 isolates, L. Weinert showed there was a complete lack of phylogeographical structuring of *S. suis* and that the genomes were very diverse, highly recombinogenic and had experienced dynamic gene gain and loss. There was strong genetic differentiation among commensal, respiratory and systemic disease isolates, although no SNP or gene was able to unambiguously resolve the three phenotypic groups. In contrast, there seemed to be little difference between human and pig isolates and, consequently, little evidence for adaptation to humans.

Furthering the topic of adaptation of bacterial pathogens to different host species, Nathalie van der Mee-Marquet (Centre Hospitalier Universitaire, Tours, France) summarized the host switching process and genomic changes of *S. aureus* clonal complex 398 (CC398). CC398 is primarily a livestock-associated clone found in pigs, but is also able to successfully infect humans and is increasingly being found in humans with no associations with livestock [47]. Based on characterization of the prophages of CC398, it was shown how these elements may play a role in host adaptation. Interestingly, the phi-MR11-like prophage appears to mediate expression of virulence genes encoded on a second prophage, thereby conferring increased invasion of human cells.

9. Round table: the global microbial identifier initiative

We are in a transition phase of clinical microbiology, with genomic sequencing being increasingly used for diagnosis

and epidemiology. There is a pressing need for tools enabling automated detection of clusters and for follow-up of epidemiological trends of clones, virulence or resistance markers. René Hendriksen (DTU, Copenhagen, Denmark) presented a talk envisioning global surveillance in the genomic era. Bioinformatics web-based tools were developed at Copenhagen's Center for Genomic Epidemiology (www.genomicepidemiology.org) to allow global epidemiology based on whole-genome sequencing. Among these tools, "ResFinder" is intended to allow the detection of acquired antimicrobial resistance genes, whereas the "PathogenFinder" and "VirulenceFinder" tools can be used to detect bacterial pathogenicity-associated genes. Automated solutions for extracting typing data, plasmid sequences, prediction of bacterial species using a k-mer approach or 16S ribosomal RNA sequences and SNP-based phylogenies were also presented. It was concluded that whole-genome sequencing combined with automated bioinformatics tools has the potential to replace many steps in conventional clinical bacteriology analyses.

John Rossen (University Medical Center Groningen, Groningen, The Netherlands) followed on with his talk on the implementation of HTS in daily clinical practice, which will include identification of strains as well as drug resistance information, epidemiologically relevant information to find related isolates and information relevant to patient management. J. Rossen described a practical example of the use of HTS for infection control based on a nosocomial outbreak of extended spectrum beta-lactamase-producing *K. pneumoniae*. HTS was used to identify candidate markers and develop PCR assays to be promptly implemented when a new case is suspected.

Dag Harmsen (Universitätsklinikum Münster, Münster, Germany) stimulated discussion about the need to reach global consensus on the implementation of genomics for the diagnosis and surveillance of infectious diseases and invited meeting participants interested in shaping the future of genomics to join the Global Microbial Identifier (GMI) initiative. GMI is a cross-sectoral, cross disciplinary initiative with more than 200 participants, which started at a meeting in Brussels (Belgium) in September 2011 (http://www.globalmicrobialidentifier.org/). René Hendriksen (DTU, Copenhagen, Denmark) detailed the main priorities of the five GMI working groups: political challenges; outreach and building a global network; repository and storage of sequence and meta-data; analytical approaches; ring trials and quality assurance; and pilot projects. During the discussion, the need for standards, on the one hand, and for curated databases containing genomic data of good quality, on the other, were particularly underlined.

10. Resistance emergence, diagnosis and epidemiology

The emergence of new resistance phenotypes challenges our ability to treat infections. Laurent Poirel (University of Fribourg, Fribourg, Switzerland and INSERM, Hospital Bicêtre, Paris, France) presented an overview of the emergence of resistance in Gram-negative bacteria and on the epidemiology of carbapenem-resistant *Enterobacteriaceae* (CRE). The necessity to evaluate the clonality of bacterial isolates from within patients up to the international level was stressed, and commonly used tools were reviewed. Although PFGE has been recognized as the "gold standard" typing method, MLST represented a very significant improvement in molecular typing, as it allowed for the definition of major emerging clones of key organisms in the group of CRE, *E. coli* and *K. pneumoniae* [8]. Since antibiotic resistance genes can be carried either on the chromosome or on mobile genetic elements such as plasmids, transposons or integrons, the epidemiology of CRE is complex. The second part of the talk illustrated key data on the international epidemiology of selected CRE, tracing their spread and dissemination [39]. The major focus lay on the clonal structure of KPC-producing *K. pneumoniae*, which harbor many different plasmids carrying a plethora of different bla_{KPC} genes. The spread of OXA-48-producing *Enterobacteriaceae* clones was also presented. On the other hand, the spread of NDM-1-producing isolates was neither linked to the dissemination of a single epidemic plasmid nor to that of specific clones.

Alessandra Carattoli (National Institute of Health, Roma, Italy) discussed the contribution of plasmids to the dissemination of antimicrobial resistance in Gram-negative bacteria. Horizontal transfer of plasmids is the main mechanism of acquisition of new resistance determinants and occurs among different species, genera and even phyla. The speaker introduced a plasmid MLST database (http://pubmlst.org/plasmid/) and free bioinformatics tools for the *in silico* detection of plasmids [29]. The use of these tools was illustrated for the IncI1 and IncF plasmid families and for the origin of dissemination of bla_{NDM-1}. Comparative analyses of plasmid structures from different CRE revealed a mosaic pattern and clearly showed the complexity of plasmid evolution.

Easy-to-use diagnostic assays are needed for public health intervention in countries with a heavy burden of MDR-TB or HIV-TB co-infection. In the next presentation entitled "High throughput combined methods for resistance emergence analysis, diagnosis and epidemiology; the multi-drug-resistant tuberculosis paradigm", Christophe Sola (Institut de Génétique et Microbiologie, Orsay, France) introduced a new technology able to provide diagnosis, molecular typing information and antimycobacterial susceptibility data for *M. tuberculosis*. The TB-SPRINT (TuBerculosis-SPoligo-Rifampin-IsoNiazid-Typing) method was developed on a high-throughput device using microbead-based flow cytometry [17]. The evaluation of the performance of this assay compared to sequencing was presented.

Artur Sabat (University Medical Center Groningen, Groningen, The Netherlands) presented his work on a whole-genome sequencing investigation of mechanisms conferring oxacillin susceptibility in *mecA*-positive *S. aureus* clinical isolates. A. Sabat characterized variation between oxacillin-sensitive (OS) and oxacillin-resistant (OR) *mecA*-positive *S. aureus* isolates [21]. An interesting difference lay in the plasmid of the OS isolate, harboring a beta-lactamase gene and

its regulatory genes. Sequence comparisons revealed that the 5′-end of one of the regulatory genes was deleted, indicating that the OS isolate would produce a truncated non-functional sensor-transducer.

11. Molecular epidemiology of viruses

In his talk entitled "Emergence and phylodynamics of rabies virus" Hervé Bourhy (Unit of Lyssavirus Dynamics and Host Adaptation, WHO Collaborating Center for Reference and Research on Rabies, Institut Pasteur, Paris, France) presented his ongoing work on rabies, a model for studies of cross-species transmission by zoonotic viruses. Despite the fact that 12 *Lyssavirus* species can infect a large range of hosts, including numerous Chiroptera species, more than 98% of human infections are due to the dog-adapted rabies virus (RABV), transmitted through dog bites. The study of nucleoproteins and glycoproteins allowed exploration of the spatiotemporal dynamics of rabies in dogs and the processes responsible for its maintenance and re-emergence. The most recent common ancestor of RABV was estimated to have circulated on the Indian subcontinent 1500 years ago. The main drivers of canid RABV phylogeography are genetic drift and population subdivision. In Africa, anthropogenic factors contributed to the spatial distribution of RABV. Moroccan and Algerian strains were clearly clustered separately, indicating that political boundaries were acting as efficient barriers to gene flow. Road distance is associated with RABV diversity, with human-mediated dispersal being inferred as a dominant factor in strain exchange between large cities. The periodic introduction of new lineages into cities leads to limited but recurrent outbreaks. Vaccination is a pivotal element for rabies control, but control of dog populations should also be implemented.

Knowledge of viral biodiversity is still sparse, especially in complex ecosystems such as the two most thoroughly studied mammalian families (*Rodentia* and *Chiroptera*). Christian Drosten (University of Bonn, Bonn, Germany) reported recent work on "Virus ecology: studying animal reservoirs to understand emergence". Through a study of insect viruses based on the partial sequence of the polymerase gene and on the spike gene, a new mosquito-borne virus named the Cavally virus (CAAV) was found which shares a common ancestor with *Coronavirinae* and *Torovirinae* [49]. C. Drosten reminded the audience that discovering new viruses in wild hosts was an important step, but did not provide many clues about emergence potential. To investigate this question, the functionality of crucial proteins of the virus was analyzed, using as a model coronaviruses detected in European bats. It was shown that the open reading frame 6 from a bat virus was able to induce the depression of the interferon response when combined with the SARS-CoV backbone in a cell-culture model [12]. Further functional studies, in particular on the diverse viruses from small mammals, are needed to better ascertain viral threats.

In late 2012, the emergence of a new *Betacoronavirus* (MERS-CoV) was reported in countries of the Arabic peninsula. To date, WHO has been informed of a total of 77 laboratory-confirmed cases of infection with MERS-CoV, including 40 deaths. Meriadeg Le Gouil (Environment and Infectious Risks Unit, Institut Pasteur, Paris, France) described the complete sequencing of the MERS-CoV genome using a combined HTS-Sanger approach and phylogenetic analyses of new strains from two cases involving nosocomial infection. Six genomes from clinical samples or virus isolates of the two MERS-CoV FRA-IPasteur cases described in France [19] were obtained. Sample preparation and HTS were depicted as pivotal for rapid exploration, directly from the clinical specimen, of such a large viral RNA genome. This strategy avoids cell culture and polymerase-induced biases. In addition, based on phylogenetic analysis of all available MERS-CoV strains, the date of emergence of MERS-CoV in humans was estimated at around July 2010. The monophyly of the human MERS-CoV was confirmed and the observed phylodynamic pattern supported a model of repeated spill-over from infected animals rather than a sustained human to human transmission model.

Muriel Vayssier-Taussat (INRA, Maisons-Alfort, France) reported on a collaborative program on the use of HTS to uncover unexpected pathogens in ticks from Western Europe. Ticks, in particular *Ixodes ricinus*, are well known for their role as vectors in the natural cycles of either bacterial or viral pathogens. However, their microbiome, including the viruses they harbor, is still largely unknown. A collection of ticks from Alsace and Ardennes regions of France were screened for pathogens by HTS. From the 174,841 assembled contigs, 12% were assigned to bacteria, 0.6% to viruses and 3% to parasites. *Bunyaviridae, Reoviridae* and *Rhabdoviridae* were among the main families represented in the pool; notably, no tick-borne encephalitis virus sequence was reported. Additionally, a Coltivirus (*Reoviridae*), the Eyach virus, was completely sequenced and isolated from mice that died following injection of tick extracts. A Nairovirus with 80% identity with the Crimean-Congo hemorrhagic fever virus (CCHFV) and a Phlebovirus (40–80% identity with the Toscana virus) were also detected. The circulation of the detected viruses in wild animals and in humans remains to be investigated.

12. Typing and surveillance networks in practice

The application of molecular typing methods to relevant public health pathogens has become deeply integrated into infection control strategies as an aid in understanding the introduction and spread of pathogens at the local, national and international levels. Clinical microbiologists and public health specialists clearly recognize the need for greater investment in whole genome sequencing. The first talk of the session entitled "Implementing whole genome sequencing: early experiences from a *Salmonella* Reference Laboratory" by Kathie Grant (PHE, Colindale, UK) offered a first insight into a switch from conventional, standardized methods to HTS, taking *Salmonella* as an example. In the face of the >1500 existing *S. enterica* serovars, serotyping as well as molecular typing methods such as PFGE are unsatisfactory due to technical issues and their limitations in recovering phylogenetic

relationships. As it was shown that MLST is a desirable alternative to serotyping [1], HTS was tested for applicability within current routine services. Clearly, this approach provides more information than conventional typing methods, revealing information about the presence of virulence factors, antimicrobial genes and SNPs. Although the speaker expressed optimism that HTS can be performed routinely in parallel with conventional typing methods, she also stressed that application and interpretation have to be extensively validated and standardized.

Marc Struelens (ECDC, Stockholm, Sweden) presented current perspectives for integrating molecular typing within European surveillance systems. In 2012, the ECDC initiated molecular typing activities covering up to 16 pathogens and five diseases. In view of the different standards and capabilities of European countries, the methods should be driven by practical needs for surveillance, they should be discriminatory in a meaningful way, they should produce comparable data across laboratories, be cost-efficient and be interpreted and assessed by experts. Collection and coordinated analysis of molecular typing data will be integrated into TESSy. Multidrug-resistant *Neisseria gonorrhoeae*, carbapenem-resistant *Enterobacteriaeceae, Acinetobacter baumannii* and multidrug-resistant *S. aureus* are among priority pathogens. Marc Struelens also elaborated on the challenges and opportunities of HTS applications to public health [44].

In the subsequent talk entitled "European molecular typing database for food, environmental and veterinary *Listeria monocytogenes* and application for the European baseline survey", Sophie Roussel (ANSES, European Reference Laboratory for *L. monocytogenes,* Maisons Alfort, France) introduced a database system for sharing and centralizing molecular data from *L. monocytogenes* isolates from food. The speaker underlined the need to strengthen and harmonize PFGE data exchange [11]. The successful integration of information into the database and, importantly, comparison with the human *L. monocytogenes* isolates, will lead to a better estimation of the importance of certain food products as the source of human listeriosis.

The importance of enterovirus surveillance was brought home by Kimberley Benschop (National Institute for Public Health and the Environment, Bilthoven, The Netherlands) during the last talk of this session, "Molecular surveillance of enterovirus in the Netherlands: evaluation of VIRO-Typened". VIRO-Typened is a web-based database tool where laboratories enter sequence-based information together with clinical and epidemiological data [27]. The usefulness of this enterovirus (EV) surveillance tool in the Netherlands and its value in the detection of clinically relevant types was evaluated in comparison to classical surveillance [38]. It offers an infrastructure on a national basis to aid poliovirus eradication and to improve EV surveillance, prevent outbreaks and provide better patient management. In the future, submission of clinical data should provide a more in-depth analysis of type-specific illnesses. Analysis of clusters of human parechoviruses was also presented, where molecular sequence data enabled analysis of possible immune divergent strains.

13. Novel diagnostic methods

Novel diagnostic tools are intended to enable better patient care and to improve epidemiological knowledge of pathogens, leading to better control of infections. In the first presentation of the session, entitled "Novel diagnostic technologies for next-gen clinical microbiology and infection control", Alex Friedrich (University Medical Center Groningen, Groningen, The Netherlands) first underlined the promise of HTS for the detection of intra-hospital transmission and presented an example of a regional health care cluster of methicillin-resistant *S. aureus* [26]. Currently, the application of HTS for infection control purposes is still limited due to a long turn-around time, with approximately four days from library preparation to phylogenetic analysis. However, in the near future, the process could take less than one day. A. Friedrich presented various diagnostic avenues and innovative technologies, including mass spectrometry, deep sequencing but also the more futuristic approaches of atom-by-atom engineering, magnetometry, electronic tables to interact with the patient or nanotechnology-based in vivo and personalized diagnoses using nanocapsules. Treatment and follow-up are also expected to evolve, with a role for diagnostic gene expression profiling based on whole transcriptome sequencing, metagenome sequencing or pharmacomicrobiomics, and 'intelligent antibiotics'. Finally, 'next-gen communication' within the health care setting should revolutionize and speed up information flow.

Matthew W. Gilmour (Diagnostic Services Manitoba, Winnipeg, Canada) discussed "Real-time genomics for bacterial pathogen surveillance". He considered the benefits and challenges represented by pathogen genomics to clinical and public health sectors. The speaker also underlined the potential of proteomics to facilitate the investigation of outbreaks. An integrated network coordinating the monitoring, assessment, investigation, outbreak response and recovery with standardized genomic analyses and unified laboratory workflows was envisioned.

In the last talk of the session, "Rapid taxonomic assignment from high throughput sequencing reads", Anthony Underwood (Bioinformatics Unit, PHE, Colindale, UK) described the bioinformatics component of the PHE project to replace existing phenotypic or molecular typing tests by HTS. *Salmonella, S. pneumoniae* and *S. aureus* were chosen in a pilot study to extract MLST, predict serotype or determine virulence gene profile. One identified limitation was the lack of accuracy in species identification and the issue of mixed cultures. Two identification methods, ribosomal MLST [23] and k-mer quantification in the sequencing reads, were described and compared. Although both methods were able to predict species correctly, the k-mer approach was faster.

14. Novel typing methods for pathogen surveillance

A large number of typing methods have been developed to assess strain relatedness for outbreak detection, surveillance, population structure and evolution studies. However, the

unsatisfied need of combining high discriminatory power, accuracy, ease of use, low cost or speed drives a continuous search for new methodologies. Repetitive DNA motifs in bacterial genomes constitute important targets for microbial typing. In the first two talks of this session, characterization of the variability of the spacer content of the clustered regularly interspaced short palindromic repeats (CRISPR) loci was presented as a valuable typing approach for two major human pathogens, *S. enterica* and group B *Streptococcus* (GBS). CRISPR loci, known to provide immunity against mobile genetic elements in some bacteria, consist of several non-contiguous direct repeats separated by stretches of variable sequences named spacers, and are often adjacent to *cas* (CRISPR-associated) genes, a set of conserved functional protein-coding genes [35]. Francois-Xavier Weill (Institut Pasteur, Paris, France) presented "CRISPR markers for *Salmonella* epidemiology" as a new method to perform real-time simultaneous typing and subtyping of all referred *Salmonella* isolates [10]. As already discussed, the current standard, serotyping, is laborious and has limitations in subtyping isolates from two highly prevalent serotypes, Typhimurium and Enteritidis. Spacer content was strongly correlated with both serotype and MLST. The study of 150 Typhimurium strains showed that spacer microevolution discriminated numerous subtypes within these prevalent serotypes. Further, a high-throughput assay for Typhimurium and its monophasic variant, the CRISPOL method, was developed based on spacer presence/absence using Luminex technology. CRISPOL is now routinely used and its database includes 450 types from 6000 isolates.

Claire Poyart (INSERM, Paris, France) showed how CRISPR spacer analysis could also be used for population studies and GBS strain typing and subtyping. The French National Reference Center for streptococci collects invasive GBS strains, which are characterized by molecular capsular typing (CPS) and MLST, but these methods are time-consuming. As was the case for *Salmonella*, *Streptococcus agalactiae* has two CRISPR loci in the genome. The CRISPR1 was sequenced in 351 GBS, non-redundant isolates previously characterized by CPS and MLST [33]. CRISPR1 trailer-end spacer content was found to be highly concordant with MLST. In contrast, the distribution of more recent spacers distinguished STs or specific lineages within the hypervirulent clonal complexes CC17 or CC23. The method was validated in a blind test with more than 100 GBS clinical isolates and was able to determine the ST of 95% of the isolates. It was concluded that CRISPR sequencing in GBS, with around 900 described spacers to date, proves to be rapid, accurate and cost-effective.

In the last talk of the session, *Salmonella* typing came back on stage, with Vitali Sintchenko (Marie Bashir Institute for Emerging Infectious Diseases & Biosecurity, Sydney, Australia) presenting "No innovation without evaluation: added value of multiple-locus variable-number-tandem-repeat analysis (MLVA) for prospective surveillance of *S. enterica* serovar Typhimurium". Surveillance of *Salmonella* infections requires high-resolution systems that detect outbreak and guide prospective interventions. A total of 8936 *S. enterica* serovar Typhimurium (STM) isolates were genotyped between January 2008 and December 2010, with the overwhelming majority of isolates belonging to only two phage types (PT), PT170 (44.6%) and PT135 (13.9%) [43]. STM isolates were further subtyped using a five-loci MLVA scheme [32], distinguishing 1562 MLVA patterns. Isolates from epidemiologically linked patients showed indistinguishable MLVA patterns. A cluster of infection was defined by $\geq$5 isolates with identical MLVA patterns obtained within a four-week period, with 50% of the isolates having been recovered in a restricted geographic area [43]. An illustration of the investigation of two outbreaks was presented. Isolates from one of these were further investigated by HTS, showing only 1 to 2 SNP differences. During the course of other outbreaks, the rise of endemic clones, in particular the STM 'super clone' ST456 cluster, was noted. This work enables more precise surveillance of *Salmonella* infections in Australia.

15. Round table on "MLST in the genomic era"

The round table was introduced by Anthony Underwood, Kathy Grant (both from PHE, Colindale, UK) and Keith Jolley (University of Oxford, Oxford, UK). It convened MLST data curators, MLST web site administrators and MLST users to discuss the issue of integration of sequence data derived from genomic sequences assembled from HTS technologies. A consensus soon emerged that data quality was not a real issue given the information by several participants that HTS-derived sequences were completely concordant with Sanger sequencing-derived MLST data. A discussion followed on the future of MLST databases. It was concluded that the long-term sustainability of MLST websites should be a concern not only for users, but also for Public Health agencies and other bodies that consider integrating MLST-derived knowledge into surveillance networks. Different options for securing funding for MLST web site development and maintenance were discussed.

16. Population dynamics, transmission networks and epidemiology

In his talk entitled 'Population dynamics, transmission networks and epidemiology', Hajo Grundmann (University Medical Center Groningen, Groningen, The Netherlands) emphasized the requirement for network approaches, such as structured surveys on a European level. These surveys should include careful consideration of the sample frame, be scalable, multicentric and tailored to the specific pathogen – preferably with a web-based analysis infrastructure. This desirable practice poses challenges in terms of extra workload, consistent sample collection and data ownership. Referring to a recent publication involving 450 hospitals in 26 countries across Europe [18], H. Grundmann described the precise temporal and spatial pattern of methicillin-sensitive *S. aureus* (MSSA)- and methicillin-resistant *S. aureus* (MRSA), which were analyzed by *spa*-typing and MLVA. By investigating this scourge of modern hospitals using network approaches to

sampling and analysis, the speaker went on to show how ecological factors at different levels of the health care cascade lead to highly modular, hierarchically distributed lineages. Successful clones are disseminated efficiently by core groups of patients and through health care institutions that are central to national health care networks, such as university hospitals.

Janina Dordel (University of Cambridge, Cambridge, UK) investigated the mobile genetic elements (MGEs) that contributed to virulence and fitness of USA300 MRSA isolates. The investigated MGEs included the Panton-Valentine leukocidin (PVL) gene-containing prophage, the pathogenicity island SaPI5 and a horizontally acquired arginine catabolic mobile element (ACME). By comparing the MGEs of isolates from disease and carriage using HTS, the authors identified different MGE profiles. The majority of the USA300 MRSA isolates harbored the SCC*mec* IVa subtype, but there was also evidence for multiple events of loss or acquisition of different subtypes. The speaker concluded that, while many prophages do not seem to play an obvious role in the virulence and fitness of USA300, dependencies between SCC*mec* subtypes, ACME, SaPI5 and the PVL-containing prophage could be observed.

Samuel Sheppard (Swansea University, Swansea, UK) described one of the first genome-wide association studies (GWAS) to be applied to bacteria [42]. Along with the increasing availability of whole genome data come significant data analysis challenges. Inspired by solutions to similar challenges in human genetics, S. Sheppard presented a method by which bacterial isolate genome sequences are divided into 30 bp reads and segregated by phenotype. The model, which takes into account the isolate phylogeny, is alignment-free and simultaneously identifies variation in the core and accessory genome. This approach enables identifying features that are significantly associated with any phenotype. These features can be mapped to a finished genome to identify phenotype-associated genes or genetic elements that are candidates for adaptation to specific niches. S. Sheppard and his colleagues applied the method to identify factors responsible for adaptation of *Campylobacter* to cattle and chickens, as host switching is an important feature in the pathogenesis of zoonotic organisms such as *Campylobacter*. A seven-gene region was found. Genes in this region were almost universally present in cattle isolates, but were frequently absent in isolates from chickens and wild birds. Three of the seven genes encoded vitamin B5 biosynthesis and isolates from cattle were better able to grow in vitamin B5-depleted media. The different diets of cattle and chickens may have provided the selective pressure leading to this trait. This method will be widely applicable to other bacterial pathogens in order to understand the genetic basis of phenotype variation.

In the final talk of this session, “Genomic epidemiology of a tuberculosis outbreak in Switzerland over 21 years”, David Stucki (University of Basel, Basel, Switzerland) revisited isolates from a historical TB outbreak in 1992 in Switzerland [13] using a combination of HTS and strain-specific SNP typing. Three *M. tuberculosis* isolates belonging to the original outbreak and five control isolates were sequenced to identify mutations unique to the outbreak strain. Two outbreak-specific mutations were incorporated into a strain-specific TaqMan genotyping assay, which was used to screen all available MTB patient isolates recovered in the Canton of Bern from 1991 to 2011 (1642 isolates). This allowed identification of 79 outbreak-derived isolates from 68 patients, differing by up to 19 SNPs. Subclusters descending from the 1992 outbreak, as well as potential superspreaders, were identified, and at least six drug-resistance-associated mutations were detected. This SNP typing approach provides a cost-effective way of identifying outbreak isolates within a large strain collection, and demonstrated the legacy of the 1992 outbreak among subsequent TB infections.

17. Molecular epidemiology from global health to one health

In his presentation “From genomics to new interventions for tuberculosis and leprosy”, Stewart Cole (Global Health Institute, Lausanne, Switzerland) reminded the audience about the high disease burden of tuberculosis and leprosy, which are characterized by a very low amount of genetic diversity encountered in the two corresponding pathogens, *M. tuberculosis* and *Mycobacterium leprae*. S. Cole illustrated how pioneering studies in tuberculosis genomics shed light on the biology, pathogenesis and epidemiology of TB. He also discussed how the promise of genomics for discovery of novel drug targets met unanticipated challenges. On the other hand, genome sequencing brought novel ways to rapidly discover cellular targets of old or redesigned drugs. Finally, S. Cole demonstrated the power of combining genomics and more conventional typing approaches such as MLVA to decipher the phylogeography of an old scourge, leprosy, and to demonstrate that populations of armadillos from America represent a reservoir of the disease bacillus.

Genomic data sharing is identified as a major requirement for the successful use of HTS in public health microbiology and global surveillance, but raises challenges linked to data access and use. Alan Hay (MRC National Institute for Medical Research, London, UK) shared his experience of the Global Initiative on Sharing All Influenza Data (GISAID). The rapid evolution of human influenza A viruses is driven by a combination of progressive sequence changes in antigenic determinants and infrequent reassortment of the genomes with those of highly distinct animal influenza viruses. A. Hay described how the Global Influenza Surveillance and Response System (and the former Global Influenza Surveillance Network) coordinated by the WHO functions as a highly responsive global cooperative mechanism that provides the basis for biannual recommendations on vaccine composition, and the early detection of and rapid response to novel influenza virus infections. Recent zoonotic infections have increased awareness of the pandemic potential of deadly influenza variants, such as H5N1 or H7N9. The strong demand for transparency of available information, and especially of genome sequence data, led to the establishment in 2008 of the GISAID EpiFlu database, in order to promote responsible sharing of all influenza virus sequences and related clinical

and epidemiological data. The EpiFlu database now includes sequences from nearly 100,000 influenza viruses and has two key founding principles. First, submitters are committed to provide data with no restrictions on their use, while retaining ownership. At the same time, users of data are committed to acknowledge the providers and make every effort to collaborate with them. A. Hay described how the resulting trust among scientists fostered international collaboration for the benefit of human as well as animal health.

18. Closing remarks and conclusions

Sylvain Brisse and Alex Friedrich expressed words of thanks to all those who ensured the success of IMMEM-10 and gave their views on the remarkable progress of the field evidenced during IMMEM-10. They also pointed out future challenges and possible extensions of the next IMMEMs to broaden their scope to include not only bacterial and viral pathogens but also parasites and fungi. It was also expressed that inclusion of more researchers from other fields such as veterinary medicine or plant pathogens could bring interesting cross-fertilization with public health microbiology. We are looking forward to IMMEM-11, which will be the opportunity to consider how much the field has evolved since the Paris meeting.

Acknowledgments

The authors would like to thank H. Fsihi for important editorial help and Richard Goering for providing material for Fig. 1.

References

[1] Achtman M, Wain J, Weill F-X, Nair S, Zhou Z, Sangal V, et al. Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. PLoS Pathog 2012;8:e1002776.

[2] Agren J, Sundstrom A, Hafstrom T, Segerman B. Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. PLoS One 2012;7:e39107.

[3] Bennett JS, Jolley KA, Earle SG, Corton C, Bentley SD, Parkhill J, et al. A genomic approach to bacterial taxonomy: an examination and proposed reclassification of species within the genus *Neisseria*. Microbiology 2012;158:1570–80.

[4] Bernard C, Brossier F, Sougakoff W, Veziris N, Frechet-Jachym M, Metivier N, et al. A surge of MDR and XDR tuberculosis in France among patients born in the Former Soviet Union. Euro Surveill 2013;18:20555.

[5] Bielejec F, Rambaut A, Suchard MA, Lemey P. SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics. Bioinformatics 2011;27:2910–2.

[6] Carrillo CD, Kruczkiewicz P, Mutschall S, Tudor A, Clark C, Taboada EN. A framework for assessing the concordance of molecular typing methods and the true strain phylogeny of Campylobacter jejuni and C. coli using draft genome sequence data. Front Cell Infect Microbiol 2012;2:57.

[7] Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. Nat Genet 2013;45:1176–82.

[8] Diancourt L, Passet V, Verhoef J, Grimont PA, Brisse S. Multilocus sequence typing of *Klebsiella pneumoniae* nosocomial isolates. J Clin Microbiol 2005;43:4178–82.

[9] Eyre DW, Cule ML, Wilson DJ, Griffiths D, Vaughan A, O'Connor L, et al. Diverse sources of *C. difficile* infection identified on whole-genome sequencing. N Engl J Med 2013;369:1195–205.

[10] Fabre L, Zhang J, Guigon G, Le Hello S, Guibert V, Accou-Demartin M, et al. CRISPR typing and subtyping for improved laboratory surveillance of Salmonella infections. PLoS One 2012;7:e36995.

[11] Felix B, Niskanen T, Vingadassalon N, Dao TT, Assere A, Lombard B, et al. Pulsed-field gel electrophoresis proficiency testing trials: toward European harmonization of the typing of food and clinical strains of *Listeria monocytogenes*. Foodborne Pathogens Dis 2013;10:873–81.

[12] Frieman M, Yount B, Heise M, Kopecky-Bromberg SA, Palese P, Baric RS. Severe acute respiratory syndrome coronavirus ORF6 antagonizes STAT1 function by sequestering nuclear import factors on the rough endoplasmic reticulum/Golgi membrane. J Virol 2007;81:9812–24.

[13] Genewein A, Telenti A, Bernasconi C, Mordasini C, Weiss S, Maurer AM, et al. Molecular approach to identifying route of transmission of tuberculosis in the community. Lancet 1993;342:841–4.

[14] Gilbert MT, Rambaut A, Wlasiuk G, Spira TJ, Pitchenik AE, Worobey M. The emergence of HIV/AIDS in the Americas and beyond. Proc Natl Acad Sci U S A 2007;104:18566–70.

[15] Glasner C, Sabat AJ, Dreisbach A, Larsen AR, Friedrich AW, Skov RL, et al. Rapid and high-resolution distinction of community-acquired and nosocomial *Staphylococcus aureus* isolates with identical pulsed-field gel electrophoresis patterns and *spa* types. Int J Med Microbiol 2013;303:70–5.

[16] Gomez-Valero L, Rusniok C, Jarraud S, Vacherie B, Rouy Z, Barbe V, et al. Extensive recombination events and horizontal gene transfer shaped the *Legionella pneumophila* genomes. BMC Genomics 2011;12:536.

[17] Gomgnimbou MK, Hernandez-Neuta I, Panaiotov S, Bachiyska E, Palomino JC, Martin A, et al. Tuberculosis-Spoligo-Rifampin-Isoniazid typing: an all-in-one assay technique for surveillance and control of multidrug-resistant tuberculosis on luminex devices. J Clin Microbiol 2013;51:3527–34.

[18] Grundmann H, Aanensen DM, van den Wijngaard CC, Spratt BG, Harmsen D, Friedrich AW. Geographic distribution of *Staphylococcus aureus* causing invasive infections in Europe: a molecular-epidemiological analysis. PLoS Med 2010;7:e1000215.

[19] Guery B, Poissy J, el Mansouf L, Sejourne C, Ettahar N, Lemaire X, et al. Clinical features and viral diagnosis of two cases of infection with Middle East Respiratory Syndrome coronavirus: a report of nosocomial transmission. Lancet 2013;381:2265–72.

[20] Guy L, Nystedt B, Toft C, Zaremba-Niedzwiedzka K, Berglund EC, Granberg F, et al. A gene transfer agent and a dynamic repertoire of secretion systems hold the keys to the explosive radiation of the emerging pathogen *Bartonella*. PLoS Genet 2013;9:e1003393.

[21] Ikonomidis A, Michail G, Vasdeki A, Labrou M, Karavasilis V, Stathopoulos C, et al. In vitro and in vivo evaluations of oxacillin efficiency against mecA-positive oxacillin-susceptible *Staphylococcus aureus*. Antimicrob Agents Chemother 2008;52:3905–8.

[22] Jolley K, Maiden M. Automated extraction of typing information for bacterial pathogens from whole genome sequence data: *Neisseria meningitidis* as an exemplar. Euro Surveill 2013;18.

[23] Jolley KA, Bliss CM, Bennett JS, Bratcher HB, Brehony C, Colles FM, et al. Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. Microbiology 2012a;158:1005–15.

[24] Jolley KA, Hill DM, Bratcher HB, Harrison OB, Feavers IM, Parkhill J, et al. Resolution of a meningococcal disease outbreak from whole-genome sequence data with rapid Web-based analysis methods. J Clin Microbiol 2012b;50:3046–53.

[25] Jolley KA, Maiden MC. BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 2010;11:595.

[26] Jurke A, Kock R, Becker K, Thole S, Hendrix R, Rossen J, et al. Reduction of the nosocomial meticillin-resistant *Staphylococcus aureus* incidence density by a region-wide search and follow-strategy in forty German hospitals of the EUREGIO, 2009 to 2011. Euro Surveill 2013;18. pii=20579.

[27] Kroneman A, Vennema H, Deforche K, v d Avoort H, Penaranda S, Oberste MS, et al. An automated genotyping tool for enteroviruses and noroviruses. J Clin Virol Off Publ Pan Am Soc Clin Virol 2011;51:121–5.
[28] Larsen AR, Goering R, Stegger M, Lindsay JA, Gould KA, Hinds J, et al. Two distinct clones of methicillin-resistant *Staphylococcus aureus* (MRSA) with the same USA300 pulsed-field gel electrophoresis profile: a potential pitfall for identification of USA300 community-associated MRSA. J Clin Microbiol 2009;47:3765–8.
[29] Larsen MV. Internet-based solutions for analysis of next-generation sequence data. J Clin Microbiol 2013;51:3162.
[30] Lemey P, Pybus OG, Wang B, Saksena NK, Salemi M, Vandamme AM. Tracing the origin and history of the HIV-2 epidemic. Proc Natl Acad Sci U S A 2003;100:6588–92.
[31] Lemey P, Suchard M, Rambaut A. Reconstructing the initial global spread of a human influenza pandemic: a Bayesian spatial-temporal model for the global spread of H1N1pdm. PLoS Curr 2009;1:RRN1031.
[32] Lindstedt BA, Vardund T, Aas L, Kapperud G. Multiple-locus variable-number tandem-repeats analysis of *Salmonella enterica* subsp. *enterica* serovar Typhimurium using PCR multiplexing and multicolor capillary electrophoresis. J Microbiol Methods 2004;59:163–72.
[33] Lopez-Sanchez MJ, Sauvage E, Da Cunha V, Clermont D, Ratsima Hariniaina E, Gonzalez-Zorn B, et al. The highly dynamic CRISPR1 system of *Streptococcus agalactiae* controls the diversity of its mobilome. Mol Microbiol 2012;85:1057–71.
[34] Maiden MC, van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA, et al. MLST revisited: the gene-by-gene approach to bacterial genomics. Nat Rev Microbiol 2013;11:728–36.
[35] Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat Rev Genet 2010;11:181–90.
[36] Mellmann A, Harmsen D, Cummings CA, Zentz EB, Leopold SR, Rico A, et al. Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. PLoS ONE 2011;6:e22751.
[37] Mooi FR, Hallander H, Wirsing von Konig CH, Hoet B, Guiso N. Epidemiological typing of *Bordetella pertussis* isolates: recommendations for a standard methodology. Eur J Clin Microbiol Infect Dis 2000;19:174–81.
[38] Niesters HG, Rossen JW, van der Avoort H, Baas D, Benschop K, Claas EC, et al. Laboratory-based surveillance in the molecular era: the TYPENED model, a joint data-sharing platform for clinical and public health laboratories. Euro Surveill 2013;18:20387.
[39] Nordmann P, Naas T, Poirel L. Global spread of Carbapenemase-producing *Enterobacteriaceae*. Emerg Infect Dis 2011;17:1791–8.
[40] Reichardt C, Chaberny IF, Kola A, Mattner F, Vonberg RP, Gastmeier P. Dramatic increase of *Clostridium difficile*-associated diarrhea in Germany: has the new strain PCR-ribotype 027 already reached us? Dtsch Med Wochenschr 2007;132:223–8.
[41] Sheppard SK, Jolley KA, Maiden MCJ. A gene-by-gene approach to bacterial population genomics: whole genome MLST of Campylobacter. Genes 2012;3(2):261–77.
[42] Sheppard SK, Didelot X, Meric G, Torralbo A, Jolley KA, Kelly DJ, et al. Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in *Campylobacter*. Proc Natl Acad Sci U S A 2013;110:11923–7.
[43] Sintchenko V, Wang Q, Howard P, Ha CW, Kardamanidis K, Musto J, et al. Improving resolution of public health surveillance for human *Salmonella enterica* serovar Typhimurium infection: 3 years of prospective multiple-locus variable-number tandem-repeat analysis (MLVA). BMC Infect Dis 2012;12:78.
[44] Struelens MJ, Brisse S. From molecular to genomic epidemiology: transforming surveillance and control of infectious diseases. Euro Surveill 2013;18:20386.
[45] Vallenet D, Belda E, Calteau A, Cruveiller S, Engelen S, Lajus A, et al. MicroScope–an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. Nucleic Acids Res 2013;41:D636–47.
[46] van Belkum A, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. Clin Microbiol Infect 2007;13(Suppl. 3):1–46.
[47] van der Mee-Marquet N, Corvaglia AR, Valentin AS, Hernandez D, Bertrand X, Girard M, et al. Analysis of prophages harbored by the human-adapted subpopulation of *Staphylococcus aureus* CC398. Infect Genet Evol 2013;18:299–308.
[48] Zhou Z, McCann A, Litrup E, Murphy R, Cormican M, Fanning S, et al. Neutral genomic microevolution of a recently emerged pathogen, *Salmonella enterica* serovar Agona. PLoS Genet 2013;9:e1003471.
[49] Zirkel F, Kurth A, Quan PL, Briese T, Ellerbrok H, Pauli G, et al. An insect nidovirus emerging from a primary tropical rainforest. mBio 2011;2:e00077–000111.

Sylvain Brisse*
Institut Pasteur, Microbial Evolutionary Genomics, Paris, France
CNRS, UMR3525, Paris, France

Carina Brehony
Department of Zoology, University of Oxford, Oxford, United Kingdom

Teresa Conceição
Laboratory of Molecular Genetics, Instituto de Tecnologia Química e Biológica (ITQB), Universidade Nova de Lisboa (UNL), Oeiras, Portugal

Meritxell Cubero
Microbiology Department, Hospital Universitari de Bellvitge-University of Barcelona-IDIBELL, Barcelona, Spain
CIBER de Enfermedades Respiratorias, ISCIII, Madrid, Spain

Corinna Glasner
Department of Medical Microbiology, University of Groningen and University Medical Center Groningen, Groningen, The Netherlands

Meriadeg Le Gouil
Institut Pasteur, Environment and Infectious Risks unit, Paris, France

Aurélie Renvoisé
AP-HP, Hôpital Pitié-Salpêtrière, Bactériologie-Hygiène, F-75013, Paris, France
Sorbonne Universités, UPMC Univ Paris 06, U1135, Centre for Immunology and Microbial Infections, team 13, F-75013, Paris, France
INSERM, U1135, Centre for Immunology and Microbial Infections, team 13, F-75013, Paris, France

Sam Sheppard
Swansea University, Swansea, United Kingdom

Lucy A. Weinert
Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom

*Corresponding author. Microbial Evolutionary Genomics, Institut Pasteur, 28 rue du Dr Roux, F-75724 Paris, France. Tel.: +33 1 40 61 36 58; fax: +33 1 45 68 87 27.
E-mail address: sylvain.brisse@pasteur.fr

Available online 31 January 2014