anses
French agency for food, environmental
and occupational health & safety

EuroReference

journal of Reference

Summer 2014 Issue No. 12

anses
French agency for food, environmental
and occupational health & safety

EuroReference

journal of Reference

Journal number 12

Summer 2014

EuroReference
journal of Reference

anses
French agency for food, environmental and occupational health & safety

Summary | Point of view | Focus | Methods | Research | Agenda

Journal No. 12
Summer 2014

Editorial

This issue of EUROREFERENCE is devoted to microbiology surveillance networks in Europe. This account is obviously not intended to be exhaustive, but to demonstrate the usefulness of these networks via a few examples.

You will find descriptions (in greater or lesser detail) of the work of several such networks, in Belgium (*AMCRA: Center of expertise on Antimicrobial Consumption and Resistance in Animals in Belgium*), Italy (*Integrated veterinary networks for the surveillance of zoonotic agents in Italy*) and France (*Purpose and overview of results of the Vigimyc Network for the epidemiological surveillance of mycoplasmoses in ruminants in France; ACTEOLab-Salmonella: a tool for monitoring salmonellae of non-human origin*), together with other European networks such as the database on *Listeria monocytogenes* set up by the EURL or the database on the genotypes of the *Mycobacterium tuberculosis* complex, serving a network that covers five continents. We have also included an article about a network involving a private company and a public university working together on the large-scale genotyping of Clustered Regularly Interspersed Short Palindromic Repeats (CRISPRs).

Lastly, the Point of View looks to the future, emphasising the importance of the IMMEM-10 Meeting in the ongoing development of concepts and techniques for the molecular characterisation of bacteria and viruses for surveillance purposes. The journal naturally keeps an eye on some highly technical problems (see the video on *Trichinella*) and regulatory issues (Methodological guide to the implementation of a process for airborne surface disinfection applied to containment areas).

We hope you enjoy this new issue!

The editorial team

Summary

EuroReference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Point of view

Genomic epidemiological typing of pathogens: feedback on the IMMEM-10 Conference

Paul M. V. Martin[1] (paul.martin@ANSES.fr) & Sylvain Brisse[2]
1. ANSES, Laboratoire de Lyon, France
2. Institut Pasteur, Paris, France

The 10th International Meeting on Microbial Epidemiological Markers (IMMEM) was held at Institut Pasteur in Paris from 2 to 5 October 2013. The contents of the scientific communications presented at the meeting were described in detail in a recent publication (Brisse *et al.,* 2014). More than 400 participants from 40 different countries attended the meeting, which included 72 oral communications, 190 posters and, of course, multiple opportunities for discussion between sessions and during coffee breaks. The above show that this conference, the 10th since the first event held in Brussels in 1987, has been a successful scientific meeting; but is was not only that: the IMMEM-10 meeting will probably be considered a turning point in the development and use of epidemiological markers for pathogenic agents in public health. It has become clear that there will be a before and an after IMMEM-10. Here are the reasons why.

If we exclude the two welcome addresses, the general introduction on challenges in public health surveillance and the outstanding tribute paid to Mark Achtman and Brian Spratt, at least 40 communications out of 67 (60%) concerned at least one of the following key words: WGS (whole genome sequencing), NGS/HTS (next generation sequencing, high-throughput sequencing), pan-genome analysis, genome comparisons, microbial genome (or whole genome) analysis, or 'genome-wide'. Several other communications implicitly referred to the whole genome sequences of studied pathogens. The place given to the use of whole genomes in public health is a true landmark that we would like to highlight here.

A number of communications discussed the sequencing of dozens or hundreds of genomes of the same bacterial species: 957 genomes of *Clostridium difficile*, 237 genomes of Shiga toxin-producing *Escherichia coli* (STEC), 25 vancomycin-resistant *Enterococcus faecium*, 111 uropathogenic *Escherichia coli*, etc. Delegates from Public Health England (PHE) presented the first results on sequences of 1500 strains of *Salmonella*: 1000 S. Typhimurium, S. Typhi, and the most commonly found serovars in 2012, and 500 strains of other serovars. The conclusion for *Salmonella* was that there is no full congruence between the serovars, the current standard epidemiological biomarkers, and the results of WGS, confirming the results already obtained by MLST (multi-locus sequence typing) (Achtman *et al.,* 2012). It would therefore be necessary to entirely rethink the current epidemiological "classification" systems used in public health, by inventing new nomenclatures. As was put somewhat provocatively by Mark Achtman, in the near future we will have to "forget our gels" and genomic epidemiology will gradually replace "fingerprinting" methods.

These results, like those obtained for other bacterial species, pose a recurring question: should the systematic use of these new methods integrate data obtained over many decades with typing methods that have become "conventional" today, such as serotyping, MLST markers, PFGE (pulsed-field gel electrophoresis) or MLVA (multiple-locus VNTR analysis)? Although they now appear to be insufficient, traditional typing methods have proven their effectiveness as microbiological tools for use in public health.

It would not therefore be desirable, from a public health decision-making standpoint, and for methodological reasons, to loose the correspondence with molecular typing data accumulated over more than a quarter of a century, and the associated epidemiological knowledge on the spatial and temporal distribution of strains and their preferential association with various sources of infection. New data from WGS are rapidly proving their value in public health, on the basis of actual experience, during outbreaks or significant events that affect pathogen population dynamics (spread of a high pathogenicity clone, of a resistance plasmid, etc.). How can we reconcile changing practices made possible and desirable by high-throughput sequencing technologies without creating a rupture with former practices, which would be damaging for decision-making in public health? There are various solutions. Two communications at the meeting showed that classic typing data can still be integrated in the era of genomic epidemiology. F.-X. Weill from the Institut Pasteur in Paris presented the use of CRISPR markers (clustered regularly interspaced palindromic repeats) and their application in *Salmonella* epidemiology. This relatively new method can be used to perform simultaneous typing and subtyping of all *Salmonella* in real time (Fabre *et al.,* 2012). Characterisation of spacer variability of CRISPR markers is today a validated typing method for *Salmonella*. The study of 150 strains of serovar Typhimurium showed that the microevolution of spacers could be used to identify and individualise many subtypes of this major serotype. Sequences or presence/absence of these spacers, identifiable through conventional methods in molecular epidemiology (including the CRISPOL method using Luminex technology), are two characteristics of strains that can easily be extracted from the genome sequence.

In the same way MLST data, used as nomenclature reference for bacterial clones can easily be deduced from genomic sequences. Keith Jolley from the University of Oxford presented the concept of gene-by-gene genomic epidemiology and the bioinformatics tool associated with the Bacterial Isolates Genome Sequence Database (BIGSdb), which extends the

Point of view

concept of the MLST method to the entire genome (Maiden *et al.,* 2013).

The system makes it possible to develop a database of bacterial strains for each pathogenic species, in which the genomic sequences and metadata associated with each strain are stored. The BIGSdb system also contains a database that defines, one by one, all the genes of the species (pan-genome). It is then possible to define any combination of genes, called schemes, useful for strain genotyping. Genotyping schemes can include different numbers of genes, for example 7 genes like in MLST schemes, or several thousands. This flexibility enables the degree of discrimination of typing schemes to be modulated based on specific needs: for example, a few dozen genes may be sufficient to identify international clonal groups, while the pan-genome may be needed to decrypt transmission events during a localised outbreak. The BIGSdb system can also be used to define schemes corresponding to groups of genes of interest (virulence, resistance). Accessible via its web interface, this system is a simple and fast tool for extracting from genomic sequences, medically important data. Moreover, this tool and other equivalent systems under development are designed to enable each community of expert microbiologists on a given pathogen to define algorithms that can be used to establish the correspondence between the genome sequences and traditional epidemiological markers (Jolley and Maiden, 2010).

These two examples show that we have entered a transition period, rather than reaching a breaking point. This transition towards pan-genomic molecular epidemiology will help microbiologists working in the area of surveillance of pathogens and outbreak control, to continue fulfilling public health requirements. It can be expected that data from whole sequences of isolates obtained during outbreaks and other important events in public health will provide new knowledge on the circulation of pathogenic agents and the epidemiology of the diseases they cause.

References

Brisse S, Brehony C, Conceicao T, Cubero M, Glasner C, Le Gouil M, Renvoise A, Sheppard S, Weinert LA. 2014. Microbial molecular markers and epidemiological surveillance in the era of high throughput sequencing: an update from the IMMEM-10 conference. Res Microbiol **165**:140-153.

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.

Achtman M, Wain J, Weill F-X, Nair S, Zhou Z, Sangal V, Krauland MG, Hale JL, Harbottle H, Uesbeck A, Dougan G, Harrison LH, Brisse S, and the S. enterica working Group. 2012. Multilocus Sequence Typing as a Replacement for Serotyping in *Salmonella* enterica. PLoS Pathog **8**:e1002776.

Maiden MC, van Rensburg MJ, Bray JE, Earle SG, Ford SA, Jolley KA, McCarthy ND. 2013. MLST revisited: the gene-by-gene approach to bacterial genomics. Nat Rev Microbiol **11**:728-736.

Jolley KA, Maiden MC. 2010. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics **11**:595.

EuroReference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Methods

A CRISPR genotyping network in France and Europe: A 5-year experiment for Research, Training and Global Health.

Christophe Sola[1] (christophe.sola@u-psud.fr), Guislaine Refrégier[1] et Michel Gomgnimbou[1,2].
1. Institut de génétique et microbiologie, équipe IGEPE et services de génotypage « Beads4Med », CNRS-université Paris-Sud, UMR8621, Orsay, France.
2. Centre Muraz, Bobo-Dioulasso, Burkina Faso.

We created a partnership and laboratory networking initiative between a private industrial company, Luminex Corporation (TX), and a public laboratory, Institut de Génétique et Microbiologie (IGM, UMR8621), for the development of high-throughput microbead-based multiplexed technologies for public health, in relation to the growing interest in polymorphic genetic objects, i.e. CRISPRs (Clustered Regularly Interspersed Short Palindromic Repeats). These loci are known to be useful for surveillance and control of infectious bacterial diseases; we show how an initial focus on a respiratory disease, tuberculosis, led us to an interest in food-borne pathogens and to new perspectives.

Keywords: CRISPR, multiplex, microbead-based hybridization, *Legionella pneumophila, Salmonella enterica, Mycobacterium tuberculosis.*

Context and History of CRISPR research; typing the *Mycobacterium tuberculosiss*

The first characterization of a CRISPR locus was made in *E. coli* by Ishino *et al.* in 1987. At that time, the CRISPR acronym did not exist; it was created in 2002 by Jansen et al. In 2004, a paper by the CDC (*Centers for Disease Control*) in Atlanta described the switch from 2D to 3D DNA chips technology for the tuberculosis "*spoligotyping*" technique, developed in 1997 by Kamerbeek et al. (Cowan *et al.* 2004). This genotyping technique assays the diversity of the CRISPR locus in the *M. tuberculosis* complex (MTBC) and remains one of the first-line genotyping techniques used for molecular epidemiology of tuberculosis. International spoligotyping database projects (SpolDB1 to SpolDB4 and now SITVITWEB[3]) helped provide high visibility for spoligotyping and enabled us to increase our knowledge concerning both the phylogeographical population structure of the MTBC and the evolution of CRISPR loci. There are now around 840 published references with *spoligotyping* as a keyword in PubMed.

MTBC, the agent of human and bovine tuberculosis, harbours one unique and "frozen" CRISPR locus, with a Direct Repeat (DR) of 36 bp with the consensus sequence: 5'-GTCGTCAGACCCAAAACCCCGAGAGGGGACGGAAAC-3'. A total of 94 spacers have been published, although only 68 were found to be specific to the MTBC and 26 were only found in the *M. canettii* ecotype. Almost no genetic variation within the DR repeat was found.

Multiplex-bead-array assays (MBAA) are considered to be a recent innovation, although single-cell analysis can be found in the literature as far back as 1977. MBAA can be used both for immunological and molecular ligands (Dunbar *et al.* 2006). Multiplexing (defined as a quantitative assay of multiple analytes simultaneously in small volumes of material) of up to 500 coloured bead types, each with its own spectral signature, as well as lasers, microfluidics and bioinformatic innovations, provide robust devices and equally robust results. The potential cost- and time-savings when compared with single-plex methods provide a strong incentive for the routine use of these methods both in research and clinical laboratories. Indeed, multiplexing technologies are to genetics what optical fibres are to information technologies: a way to have multiple information channels in a single pipe and to produce and transmit a deluge of data. This amazing process should however of course be complemented by data storage and data analysis processes (e.g. by cloud computing and data-mining systems).

When creating the IGEPE team in 2007, the business development unit of Luminex in the Netherlands (Luminex BV, Oosterhout) helped us establish microbead-based flow cytometry in our research team. The original team actually consisted of only two people, but grew rapidly. No laboratory in Europe had tried to reproduce the transfer of spoligotyping in bead-based format; this first step had been done at the CDC as previously mentioned, and Zhang *et al.* extended in our laboratory the original 43-spacer spoligotyping format to a 68-spacer format, which was shown later to be more efficient for achieving better discrimination in M. bovis and South-East-Asian genotypes of MTBC clinical isolates (Zhang *et al.* 2010, 2011).

With funding from the "Région Ile-de-France" and technical assistance from Luminex BV, we started to provide genotyping services, training, and sales of microbead-based spoligotyping services, and were also able to develop innovative non-catalogue coupled microbead-based assays with the support of a few customers in Europe (the RIVM in the Netherlands) and especially in France (Inserm research team in Montpellier, Bichat Hospital and National Reference Tuberculosis Centre in Paris, ANSES teams). We trained dedicated staff within these teams and looked at new ways of applying microbead-based techniques to public health issues. **Figure 1** shows a brief chronology of method development over the last six years.

3. Spoligo-International Type, VNTR-International Type. http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/.

Methods

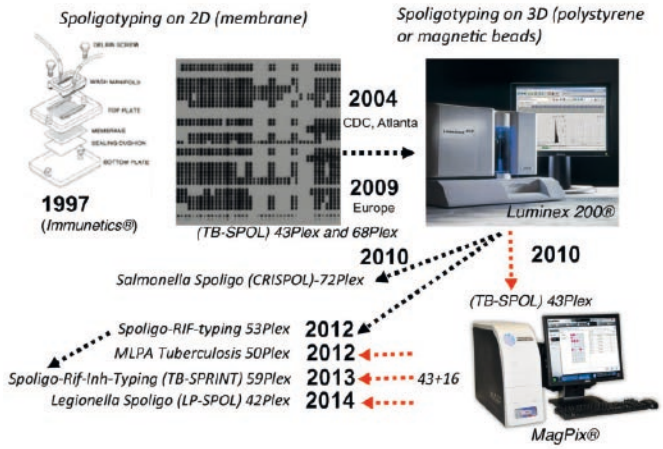


Figure 1. Brief chronology of the development of microbead-based methods on Luminex platforms; since 2010, all methods have been developed on both the Luminex 200® and on the MagPix®. Black arrows: assays developed on the Luminex 200®; red arrows: assays developed on the MagPix®.

The CRISPOL technique: Application of CRISPR technique to *Salmonella enterica*

Among enterobacteriae, *Salmonella* enterica is a major food-borne pathogen and is consequently of great economic interest. For the time being, the identification of thousands of serovars remains the gold standard though it is a tedious and costly method. The initial serovar typing scheme goes back to 1934, but has been regularly updated, most recently in 2007. Even if molecular genotyping methods (PFGE, MLST, MLVA), databases and networks (PulseNet) are well implemented, there is a need for more cost-effective methods. The recent *E. coli* 0104:H4 outbreak in Germany and France also pointed up the need for an increased use of very discriminatory loci such as CRISPR loci when responding to health crises.

An exhaustive analysis of the genetic diversity of CRISPR loci in *Salmonella* was performed thanks to a huge CRISPR loci sequencing operation on the historical collections of the Enterobacteria laboratory at the Pasteur Institute, with the firm commitment of its current director, Dr F.X. Weill, and the support of the genomic and public health departments.

In *Salmonella*, there are two CRISPR loci, CRISPR1 and CRISPR2. The CRISPR1 locus is located downstream from the *iap* gene, whereas CRISPR2 is located upstream from the *ygcF* gene. The DRs of both CRISPR loci are 29 bp long and have the consensus sequence 5-CGGTTTATCCCCGCTGGCGCGGGGAACAC-3. CRISPR analysis by PCR and sequencing of 783 strains belonging to 130 serotypes revealed the presence of 3,800 spacers with a mean size of 32 bp (Fabre *et al.* 2012). The spacer content was found to be correlated with both serotype and multilocus sequence type (MLST). Furthermore, spacer microevolution (duplication, triplication, loss or gain of spacers, presence of SNP variant spacers or VNTR variant spacers) discriminated between subtypes within prevalent serotypes such as Typhimurium (STM), the most prevalent serotype worldwide. In eight genomes and 150 strains of serotype Typhimurium and its monophasic 1,4,[5],12:i:- variant, 57 CRISPR1, 62 CRISPR2 alleles and 83 CRISPR1-CRISPR2 combined alleles were found. Forty unique spacers (including four with variants, such as SNP or VNTR variants) were identified in CRISPR1. Thirty-nine unique spacers (including two with a SNP variant) were identified in CRISPR2. Particular well-characterized populations, such as multidrug-resistant DT104 isolates, African MDR ST313 isolates, and DT2 isolates from pigeons, each had typical CRISPR alleles. Based on this high polymorphism of the spacer contents, a microbead-based liquid hybridization assay, CRISPOL (for CRISPR polymorphism) has been developed. This assay targets 72 spacers identified previously.

The CRISPOL assay was patented by the Institut Pasteur, and implemented in routine both at the Pasteur Institute and at the ANSES laboratory (Fabre *et al.* 2012). It enables the identification of emerging *Salmonella enterica* serovar Typhimurium outbreaks almost in real time, using a new CT (CRISPOL type) clone nomenclature. Our team contributed to the development, implementation and optimization of this assay, and to the optimization of the reagent production process, and is now a supplier of quality-controlled coupled-microbeads to these public health laboratories. **Figure 2** describes the principles and schematizes the CRISPOL technique.

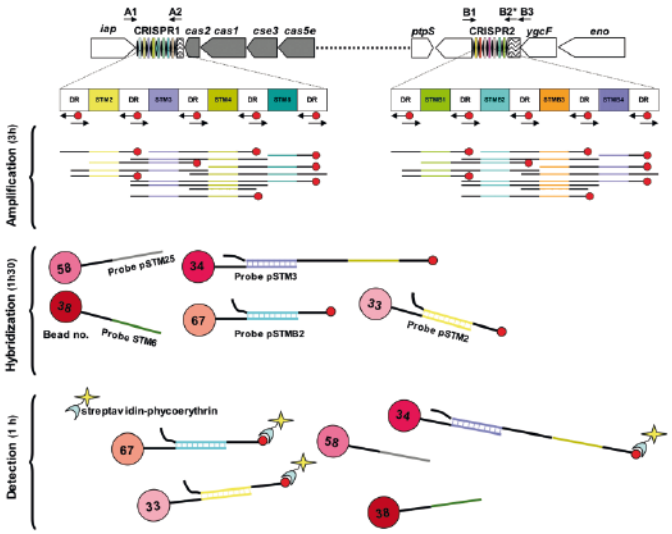


Figure 2. Description of all steps of the CRISPOL technique. Contrary to MTBC, *Salmonella* has two CRISPR loci. A total of 72 spacer probes (including 4 SNP variants) are assessed during this four-five hour method. The technique was implemented in routine analysis and can detect outbreaks in the early stages for a few Euros per assay. (Reproduced from Fabre *et al.* 2012 with authorization.)

1. Pulsed Field Gel Electrophoresis
2. Multi-Locus sequence Typing
3. Multi-Locus VNTR Analysis

Euro Reference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Methods

Legionella pneumophila ST1/pulsotype Paris and the LP-SPOL technique

Legionella pneumophila is a gram-negative facultative intracellular pathogenic bacterium that was identified as the infectious agent of Legionnaire's Disease (LD) or Legionellosis in 1977. Several species of the *Legionella* genus can cause LD, but *L. pneumophila* is responsible for the major outbreaks of LD and for more than ~90% of all identified clinical cases, and *L. pneumophila* serogroup 1 accounts for ~85%. The organism is quite ubiquitous in water environment, whether natural or artificial water networks. Its pathogenicity is related to pulmonary infections, mainly acute pneumonia, which can be fatal in quite a large number of cases (29 fatal in 182 cases in the first reported outbreak).

The reference method for studying the molecular epidemiology of *L. pneumophila* is restriction enzyme analysis using pulse-field gel electrophoresis (PFGE). This method, although laborious, remains a reference even though sequence-based typing (SBT), monoclonal antibody typing and more recently MLVA (Multiple-locus Variable Number of Tandem Repeat Analysis) have been developed. However such methods have limitations in molecular epidemiological investigations due to their lack of power of discrimination of certain strains and they are tedious and slow.

In a previous study, Ginevra *et al.* showed that in certain cases, some indistinguishable strains either by the STB or PFGE methods, particularly within the *L. pneumophila* ST1/Paris pulsotype, could be studied more efficiently using CRISPR, and they developed a membrane-based spoligotyping method. The purpose of our recent collaboration with the National Reference Centre on *Legionella* was to transfer the previously-developed membrane-based spoligotyping technique to the microbead-based format. The *Legionella* direct repeat is a 37 bp repeat with the following sequence:
5'-CCAATAATCCCTCATCTAAAAATCCAACCACTGAAAC-3'.
We recently successfully transferred the membrane-based method to microbeads, both on Luminex 200® and on a Magpix® device and are planning to launch an international multicentric study. The method appears promising for tracking certain specific clones of *L. pneumophila* (Gomgnimbou *et al.* 2014).

More sophisticated and integrated methods for diagnosis and surveillance can be developed

In 2009-2010, our team became the French beta-tester site of a new fluorescence imaging technology (MagPix®) that eliminates most of the microfluidics systems as well as the Laser technology used in flow cytometry. This technology uses a magnet that attracts spreaded paramagnetic beads on a surface, LED technology and a CCD camera, thus providing a portable, bench or field device, with a reduced 50-plex format and a slightly longer read-out time. In collaboration with other teams (the Royal Tropical Institute in Amsterdam or KIT) or by using in-house expertise, we developed more complex assays that use SNPs and deletion typing to work on this new platform, as well as other molecular biology strategies to optimize the multiplexed assays. One first solution to facilitate PCR multiplexing is to use the MLPA (Multiple Ligation-dependent Probe Amplification) principle. MLPA is used very successfully in human genetics and was applied to develop an assay on the MagPix, which simultaneously (i) identifies MTBC species (MTBC versus non-tuberculosis mycobacteria), (ii) identifies drug-resistance genotypes, and (iii) allows sub-species and even clade-level identification of genotypes based on characteristics-SNPs and signatures (Bergval *et al.* 2012). A second solution for performing such assays uses the DPO principle (dual-priming oligonucleotides) to simultaneously perform (i) spoligotyping and obtain a rifampin genetic susceptibility profile of the most frequent drug-resistance mutations or to simultaneously perform (ii) spoligotyping, obtain a rifampin and isoniazid genetic susceptibility profile on the most frequent mutations, with 90% sensitivity and 100% specificity (Gomgnimbou *et al.* 2012, 2013). The first method uses the direct assessment on positions 516, 526 and 531 of the rpoB 81 bp Hot-Spot and indirect assessment on the other positions of the rpoB gene due to "sloppy" molecular probes. The second method adds the detection of katG 315 and inhA -7 and -15 mutations, all mutations responsible for isoniazid resistance. With this assay we can thus simultaneously detect and track multi-drug-resistant tuberculosis bacilli transmission. This last assay (TB-SPRINT) is currently being evaluated directly on biological samples. Such an assay, if disseminated to well-trained reference laboratories, could also change the way drug-resistance surveys are regularly performed in developing countries with high TB prevalence, with greater efficiency and lower costs.

Outlook for CRISPR typing on new models or new technologies

Spoligotyping could soon become a generic method, since polymorphic CRISPR regions are described at an increased pace and for more and more pathogens such as *Corynebacterium diphteriae, Lactobacilli* and *Streptococci*. Spoligotyping is widely accepted internationally as a first-line method for genotyping MTBC for molecular epidemiological studies and for enhancing our understanding of the global phylogeographical structure of the *M. tuberculosis* bacilli population. When associated with drug-resistance gene mutations, it now provides a unique laboratory method for tracking the spread of MDR-TB, easier to analyze than Whole-Genome Sequencing and Next-Generation Sequencing (WGS/NGS). The possibility of running spoligotyping on flow cytometers or fluorescence imaging devices offers a number of advantages. These microbead-based systems are high-throughput, thus allowing faster flows and a better standardization of assays, making it easier to develop and implement these techniques. These systems are cost-effective, and well suited for use in routine analysis and for the surveillance and control of infectious diseases. Such typing systems, which can detect variations in CRISPR content, are dependent on the rate of spacer turnover and/or acquisition, a process that is strongly dependent on phage challenges and hence on environmental conditions.

It is too early to predict whether CRISPR-based methods will gain wide acceptance at a time when WGS/NGS and Mass Spectrometry are tending to become the preferences of clinical and public health microbiologists. It is certain that the price *per sample*, the possibility of fully automating technological platforms, and the user-friendliness of data processing and results management software applications will play a key role in the success or failure of CRISPR-based methods. Quality

Methods

controlled, niche assays (non-Luminex catalogue), that may only be run once or twice a year for very specific purposes on these devices, are unlikely to find a market unless an adequate ecosystem of many multiplexed techniques is specifically developed. This is the mission and vision of the IGEPE team and of "Beads4Med" genotyping services, to serve public health laboratories by expanding our activities through the sustained, partner-based, realistic but ambitious development of our products and services portfolio for European and world-wide public health issues.

Acknowledgments

We are grateful to the Centre national de la recherche scientifique (CNRS), université Paris-Sud, fondation Mérieux, and the région Ile-de-France for financial support, and to the Luminex BV company (M. F. Topin and M. J. van Gils) for technical support. This work also required the scientific collaboration/partnerships of several research teams, among which the following are specifically acknowledged: Dr. F.X. Weill and Dr. S. Le Hello (institut Pasteur, Paris), Dr. C. Ginevra and Dr. S. Jarraud (French National Reference Centre for *Salmonella*, Lyon), Dr. M.L. Boschiroli and Dr. A. Brisabois (ANSES, Maisons-Alfort, France). We also thank all our faithful international partners, customers and collaborators.

References

Bergval I, Sengstake S, Brankova N, Levterova V, Abadia E, Tadumaze N, Bablishvili N, Akhalaia M, Tuin K, Schuitema A, Panaiotov S, Bachiyska E, Kantardjiev T, de Zwaan R, Schurch A, van Soolingen D, van't Hoog A, Cobelens F, Aspindzelashvili R, Sola C, Klatser P, Anthony R. 2012. Combined Species Identification, Genotyping, and Drug Resistance Detection of *Mycobacterium tuberculosis* Cultures by MLPA on a Bead-Based Array. *PLoS One*, 7:e43240.

Cowan LS, Diem L, Brake MC, Crawford JT. 2004. Transfer of a *Mycobacterium tuberculosis* genotyping method, Spoligotyping, from a reverse line-blot hybridization, membrane-based assay to the Luminex multianalyte profiling system. *J Clin Microbiol*, 42:474-477.

Dunbar SA. 2006. Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. *Clin Chim Acta*, 363:71-82.

Fabre L, Zhang J, Guigon G, Le Hello S, Guibert V, Accou-Demartin M, de Romans S, Lim C, Roux C, Passet V, Diancourt L, Guibourdenche M, Issenhuth-Jeanjean S, Achtman M, Brisse S, Sola C, Weill FX. 2012. CRISPR Typing and Subtyping for Improved Laboratory Surveillance of *Salmonella* Infections. *PLoS One*, 7:e36995.

Ginevra C, Jacotin N, Diancourt L, Guigon G, Arquilliere R, Meugnier H, Descours G, Vandenesch F, Etienne J, Lina G, Caro V, Jarraud S. 2012. *Legionella pneumophila* sequence type 1/Paris pulsotype subtyping by spoligotyping. *J Clin Microbiol*, 50:696-701.

Gomgnimbou MK, Abadia E, Zhang J, Refregier G, Panaiotov S, Bachiyska E, Sola C. 2012. "Spoligoriftyping" a DPO-based direct-hybridization assay for TB control on a multianalyte microbead-based hybridization system. *J Clin Microbiol*, 50:3172-3179.

Gomgnimbou MK, Hernandez-Neuta I, Panaiotov S, Bachiyska E, Palomino JC, Martin A, Del Portillo P, Refrégier G, Sola C. 2013. "TB-SPRINT: TUBERCULOSIS-SPOLIGO-RIFAMPIN-ISONIAZID TYPING"; an All-in-One assay technique for surveillance and control of multi-drug resistant tuberculosis on Luminex® devices. *J Clin Microbiol*, 51:3527-3534.

Gomgnimbou MK, Ginevra C, Peron-Cane C, Versapuech M, Refrégier G, Jacotin N, Sola C, Jarraud S. 2014. Validation of a microbead-based format for spoligotyping of Legionella pneumophila *J Clin Microbiol*. Apr 23. [Epub ahead of print].

Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. 1987. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol*, 169:5429-5433.

Jansen R, van Embden JD, Gaastra W, Schouls LM. 2002. Identification of a novel family of sequence repeats among prokaryotes. *Genomics*, 6:23-33.

Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, van Embden J. 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol*, 35:907-914.

Zhang J, Abadia E, Refregier G, Tafaj S, Boschiroli ML, Guillard B, Andremont A, Ruimy R, Sola C. 2010. *Mycobacterium tuberculosis* complex CRISPR genotyping: improving efficiency, throughput and discriminative power of 'spoligotyping' with new spacers and a microbead-based hybridization assay. *J Med Microbiol*, 59:285-294.

Zhang J, Heng S, Le Moullec S, Refregier G, Gicquel B, Sola C, Guillard B. 2011. A first assessment of the genetic diversity of *Mycobacterium tuberculosis* complex in Cambodia. *BMC Infect Dis*, 11:42.

EuroReference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Methods

Video illustrating the artificial digestion method for Trichinella larvae detection in meat

Pauline Macé[1] (pauline.mace@ANSES.fr) ; Bruno Allouche[2] ; Aurélie Chevillot-Grasset[1] ; Aurélie Heckmann[1] ; Bruno Polack[1] ; Sandrine Lacour[1]; Pascal Boireau[1] ; Isabelle Vallée[1].

1. UMR BIPAR ANSES – Enva – Upec ; ANSES, Maisons-Alfort Laboratory for Animal Health, France, laboratoire national de référence « Parasites transmis par les aliments »,Maisons-Alfort, France

2. Ecole nationale vétérinaire d'Alfort, France

Keywords: ***Trichinella*,**

Introduction

It is estimated that approximately 60% of emergent or re-emergent human diseases are zoonoses[1] (diseases which animals can transmit to humans and vice versa). Due to the widespread distribution of affected animal species[2,3], effective prevention methods, reliable controls and heightened surveillance of zoonotic diseases are essential. One zoonotic disease, trichinellosis, constitutes a major public health issue worldwide since there is no curative treatment for it. The availability of an effective system for detecting trichinellosis-positive carcasses prior to consumption is therefore crucial.

Trichinellosis

Trichinellosis is a globally distributed zoonosis caused by consumption of raw or undercooked meat infected with *Trichinella* spp. muscle larvae. This pathogen is a nematode (roundworm) parasite with an infectious stage located in the striated muscle cells of its host (Figure 1). *Trichinella* is able to infect all monogastric mammals (in particular pigs, wild boars and horses) and some raptors or detritivore birds and reptiles.

Figure 1 : *Trichinella spiralis* muscle larvae

http://vimeo.com/user16014309/methodetrichinella
« Illustration d'une méthode de détection de *Trichinella* dans les viandes»
«Implementation and critical points of a *Trichinella* detection method in meat : artificial digestion and microscopic examination (reference method according to the UE regulation n° 2075/2005).»

Worldwide, the number of people contaminated by *Trichinella* is estimated at 11 million[4] and each year approximately 10 000 new infections occur[5].
Trichinellosis is an asymptomatic disease in animals, but in humans it is characterised by a painful presentation (diarrhoea, fever, facial oedema, muscle pain and nervous signs) for which there is no effective curative treatment. Contamination by *Trichinella* larvae can lead to sometimes irreversible sequelae and in very rare cases to death. Thus, inspection of carcasses for *Trichinella* at the slaughterhouse is a key regulatory measure for collective control. In France, these analyses are performed on carcasses of species susceptible to *Trichinella* infection (pigs, wild boars, horses) by official laboratories approved by the Ministry of Agriculture.

1. Cutler SJ, Fooks AR, van der Poel WH. 2010, Public health threat of new, reemerging, and neglected zoonoses in the industrialized world. Emerg Infect Dis, 16,1-7
2. Merianos, A., 2007, Surveillance and response to disease emergence. Curr Top Microbiol Immunol 315, 477-509.
3. Murphy, F.A., 2008, Emerging zoonoses: the challenge for public health and biodefense. Prev Vet Med 86, 216-223.
4. Dupouy-Camet, J., 2000, Trichinellosis: a worldwide zoonosis. Vet Parasitol 93, 191–200.
5. Dorny P, Praet N, Deckers N, Gabriel S. 2009, Emerging food-borne parasites. Vet Parasitol, 163, 196-206.
6. Soulé, C., Dupouy-Camet, J., 1991, Conséquences physiopathologiques et immunologiques de l'infestation par Trichinella. In: « La trichinellose: une zoonose en évolution » (ed) Office International des Epizooties, 18.
7. Pozio, E., 2001, New patterns of Trichinella infection. Vet Parasitol 98, 133-148

Methods

The detection method in meat

The reference method on the date of online publication of this video (June 2014) is the artificial digestion of blended muscle samples. This method is described in Commission Regulation (EC) No 2075/2005 of 5 December 2005 (Annex I, Chapter I), amended by Regulation (EC) No 1245/2007 dated 24 October 2007. It consists in artificial digestion of muscle samples in order to release and then detect by microscopy any *Trichinella* larvae present. Although the test's principle is simple, its implementation is fully manual and cannot be automated. The performance of the method therefore mainly relies on the technical expertise of the analyst. Furthermore, there is no possibility of including any internal positive or negative controls during the test. Maintaining a high level of technical expertise, especially by training analysts to master the critical points of the method, is therefore essential to ensure the quality of results.

A new tool to supplement the training of Trichinellosis screening analysts

To this end, the French National Reference Laboratory for Foodborne Parasites (ANSES, Maisons-Alfort Laboratory for Animal Health, France) has created a video to present a technical operating procedure for the artificial digestion method, highlighting critical points. This video can specifically be used as one of the tools for primary and continuous technical training of analysts performing the artificial digestion method.

1. This video is intended only as an illustration of an operating procedure for detecting Trichinella muscle larvae by artificial digestion in accordance with the reference method described in Commission Regulation (EC) No 2075/2005 of 5 December 2005 (Annex I, Chapter I), amended by Regulation (EC) No 1245/2007 dated 24 October 2007. This video has no legislative or regulatory value; laboratories must refer to current regulations to obtain the text for the method to be applied. ANSES shall in no way be held liable for damages of any kind for a laboratory resulting from analyses performed as per the method described in this video, particularly if, in doing so, the laboratory deviated from rules applicable on the date of the analysis.

2. This video is the property of ANSES. It is protected by international conventions on industrial, literary and artistic property rights currently in force, and by legislative and regulatory provisions in effect in France. Reusing this video is subject to compliance with the rules stipulated in the "Legal Information" section of the ANSES website

Euro Reference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Focus

Methodological guide to the implementation of a process for airborne surface disinfection applied to containment areas

P. Maris, S. Allix, S. Etienne, B. Garin Bastuji, B. Gassilloud, G. Lecarrou, N. Madani, P. Marianneau, E. Monchâtre-Leroy, F. Rizzo, E. Rousset, K. Sidi-Boumedine et S. Zini. French Agency for Food, Environmental and Occupational Health & Safety (ANSES), Maisons-Alfort, France

Introduction

Microbiology laboratories are nowadays subject to strict security constraints designed to protect workers, the public and the environment. In particular, public health concerns related to the toxicity of formaldehyde, used as an airborne surface-disinfection (ASD) agent, have led to the development of alternative processes whose use and efficacy are not always well controlled. Knowledge of the basic principles of ASD and broad user experience are of crucial importance for maintaining biosafety in laboratories. Consequently, ANSES's Committee for the Control of Laboratory Biorisks (CMRBL) decided to compile a guide on all the information available in this area that could help the user select an ASD process. Note that the disinfection techniques covered in this guide relate exclusively to automatic disinfection process , without human presence. This guide reviews the key principles of ASD use, presents the main applicable regulations and standards relating to it, and provides users with practical advice for implementing ASD.

General information on airborne surface-disinfection processes

With regard to the aims of this guide, ASD can be defined as "an operation with a temporary result that reduces contamination of inert environments or surfaces by micro-organisms to an acceptable level, depending on the objectives set by the risk assessment".

This operation is performed by means of an automatic process, whose active ingredients are chemical agents in gaseous or mist form. It is intended for disinfecting surfaces in a given volume, regardless of their orientation. Because of the toxicity of the disinfectants used, this type of process is essentially implemented in a laboratory, strictly in the absence of any human presence. This operation, which is especially vital to the operation of Biosafety Level 3 (BSL3) laboratories, is justified prior to periodic maintenance of the room, and also prior to moving a device outside the containment area, before *in situ* maintenance of a contaminated device or system, or after accidental release of infectious material.

European regulations

Products used for ASD are defined as "biocidal" products under the European Biocides Directive, 98/8/EC. The European Union has established a regulatory framework for the marketing of biocidal products to ensure a high level of protection for humans, animals and the environment. Since 1 September 2013, implementation of this Directive has been subject to Regulation (EU) no. 528/2012 of 22 May 2012, which consists of two steps:

- an assessment of biocidal active substances, that may or may not result in their inclusion on a European positive list;
- an assessment of products that contain the active substance(s), with a view to obtaining national or European marketing authorisation (MA) that meets common requirements at European level.

Among the 22 product types (PT) covered by the European regulation, in the scope targeted by this regulation, biocidal products used for ASD fall into category PT2. Full implementation of this regulation involves a transitional period that currently requires manufacturers to declare their products and processes to the Ministry of Ecology.

French regulations

The above-mentioned Biocides Directive has been transposed into French law in Articles R 522-1 to 522-47 of the French Environment Code. In addition, the Decree of 19 May 2004 defines the conditions relating to control of the marketing of biocidal active substances and marketing authorisation for biocidal products. The French National Agency for Medicines and Health Products Safety (ANSM) has been entrusted with monitoring the market for ASD processes in accordance with Article L.3114-1 of the French Public Health Code (CSP), which stipulates that "when it is necessary due to either the transmissible nature of infections of people being accommodated, treated or transported, or the risk factors for acquiring infections by people admitted to these premises or transported in these vehicles, biocidal products must be used to disinfect: 1) premises receiving or accommodating patients and those where medical, paramedical or veterinary treatments are given; 2) vehicles used for medical transport or for transporting bodies; 3) premises and vehicles exposed to the micro-organisms and toxins mentioned in Article L.5139-1 of the Public Health Code (Decree of 30 April 2012 establishing the list of micro-organisms and toxins)."

Since 2007, the French government has regularly published decrees leading to bans on biocidal substances. These bans must be linked to a given product type, and result in the suppliers withdrawing these products for a specific use. In fact, for the ASD processes covered by the ANSM's market surveillance, these bans now only concern the hospital sector. In addition, the user must choose from the products and processes that are on the market according to the specificities and constraints of the laboratory(ies) for which they are intended.

European and French standards

With regard to the claims declared by the manufacturers, in terms of efficacy, each product granted an MA must meet the requirements described in some or all of the following standards, taking into account, where appropriate, additional

Focus

requirements according to their specificities. Examples include the following standards:
- NF EN 14348 for mycobactericidal activity;
- NF EN 14476 and NF EN 14675 for virucidal activity;
- EN 1650 and EN 1657 for fungicidal activity;
- EN 1276 and EN 1656 for bactericidal activity;

It should be noted that this list is not exhaustive and is constantly evolving. The Afnor website can be consulted for updated information. For the ASD processes targeted by this guide, the French standard NF T 72-281 (2009) proposes a method for determining bactericidal, fungicidal, yeasticidal and sporicidal activity for ASD processes. It applies to automatic and manual processes, non-pressurised (spray type) or pressurised (limited to 10 bars). In early 2011, this method was proposed at European level and discussions are underway on a future European standardisation by the European Committee for Standardization (CEN/TC 216: "Methods of airborne disinfection of surfaces - Determination of bactericidal, fungicidal, yeasticidal, sporicidal and virucidal activity"). As part of the next revision of the French NF T 72-281 Standard, a chapter on the determination of virucidal activity will be included.

Focus on the use of formaldehyde and its derivatives in ASD with regard to the regulations

The provisions of Decree No. 2001-97 of 1 February 2001 establishing the specific rules for the prevention of carcinogenic, mutagenic or reprotoxic (CMR) risks and amending the French Labour Code apply to formaldehyde and any preparation containing more than 0.1% of formaldehyde. The Decree of 13 July 2006 (amending the Decree of 5 January 1993 establishing the list of carcinogenic substances, preparations and processes within the meaning of the second paragraph of Article R.231-56 of the French Labour Code) includes formaldehyde. Although formaldehyde is still authorised under PT2, this French (and perhaps future European) position led to work being reinitiated on seeking alternatives to formaldehyde. In fact, formaldehyde is classified as a Group 1 carcinogen by the International Agency for Research on Cancer (IARC) and as a Category 3 carcinogen with the risk phrase R40 (*"Limited evidence of a carcinogenic effect"*) in the European classification.

Principles of airborne surface disinfection

ASD involves applying a biocidal product to surfaces, using air as the diffusion vector. The aim of this method is to disinfect surfaces (equipment, walls, floors) by emitting a biocidal product into the atmosphere using an automated dispersion device. It is important to emphasise that this process only applies to the disinfection of surfaces and cannot under any circumstances be applied to directly disinfect the air. The device should diffuse the biocidal product in such a way that it comes into contact with all the surfaces in the room to be disinfected. At least three types of dispersion device are currently available on the market and are based on the following processes:
- Nebulisation: droplet size is less than 5 *µm*;
- Spraying: droplet size ranges from 10 to 50 *µm*;
- Flash evaporation: the heated biocidal product (e.g. hydrogen peroxide) vaporises and is drawn by an airstream into the room to be disinfected.

The level of efficacy of the biocidal products will depend on the diffusion process that is selected. Consequently, the efficacy assessment, as well as the laboratory validation of a process, apply only to an inseparable "device/product" combination.

Preparation of the room

The preparation of the room to be disinfected is an important preliminary step that should not be neglected. Inadequate preparation could otherwise lead to non-compliant disinfection results or degradation of equipment. This phase includes:
- cleaning and bio-cleaning;
- protecting sensitive devices;
- opening doors and drawers in furniture;
- installing disinfection equipment;
- verifying that it is in working order;
- configuring the air handling unit (AHU);
- verifying that the room is sealed;
- positioning biological (BIs) and chemical indicators.

The various phases of ASD

An ASD cycle can theoretically be broken down into four successive phases that can be adapted as needed:
- **the pre-treatment phase** during which the correct environmental conditions (temperature and humidity) are obtained in the room to be disinfected, in order to optimise the efficacy of the treatment. This optional phase is dependent on the chosen process;
- **the phase of dispersion** of the disinfectant by the device in the room to be disinfected;
- **the phase of contact** between the product and the surfaces to be disinfected;
- **the aeration phase** intended to remove the disinfectant before operators can re-enter the room. Environmental verifications may be considered in order to better determine how long the room will be unavailable, for ensuring safety of personnel. The maximum exposure limits (MEL) for each product must be known.

Dispersion time, contact time and particle size of the disinfectant droplets

The dispersion and contact times for the biocidal product with the surfaces to be disinfected are parameters to be considered closely:
- dispersion time is the period needed to reach the target concentration of the product on the surface to be disinfected, in a given volume;
- the contact time of the product with the surfaces to be disinfected is the duration needed to achieve the expected biocidal efficacy.

The automated devices diffuse the biocidal products either as a gas or in the form of microdroplets. Concerning the size of these microdroplets, a relationship has been established with their settling times. The example given in **Table 1** concerns a room left undisturbed and clearly demonstrates that the smaller the droplets' diameter, the greater their stability in the air.

Table 1. Relationship between biocide droplet size and settling time

Droplet diameter	0,5 µm	1 µm	3 µm	10 µm	100 µm
Settling time	41 h	12 h	1.5 h	8 min	5 sec

It is also important to note that the particle size of the droplets is dependent on the viscosity of the products. Thus, for the same process, two products of different viscosity will produce droplets of different size. Finally, the efficacy of ASD is dependent on the stability in air of the biocidal product, which, in order to maintain its biocidal power, must remain chemically

Focus

stable for a sufficiently long period. A trade-off must be found between the product's chemical stability to obtain optimal biocidal activity and its rapid post-disinfection degradation to prevent it persisting on surfaces, which is both hazardous to personnel and harmful to facilities. Therefore it is important to obtain precise information from the supplier of the process (device/product combination) about the size of the droplets, their propulsion speed in the air, or the maximum distance reached by these droplets under normal operation in a room left undisturbed, as well as the duration of the product's chemical stability.

Compatibility of the disinfectant with the materials

In a laboratory, the facilities and equipment are composed of a wide variety of materials. In addition, new equipment can be regularly introduced and new layouts arranged. It is important to ensure that the surfaces to be disinfected are compatible with the selected product. Therefore, it is necessary to obtain information from the manufacturer about the product's corrosive power and its compatibility with the target surfaces to be disinfected. This is because many of the available products are highly acidic and/or oxidising (**Table 2**). Given these characteristics, users should be aware that excessive condensation promotes the corrosion of many materials.

Authorisation of personnel and maintenance of the disinfection process

Laboratory operators need to be authorised to use, maintain and verify the disinfection process. Regarding the process, before it can be implemented, it is important to obtain precise information on its maintenance, upkeep and appropriate verifications. On some types of complex equipment, it may be necessary to establish a maintenance contract.

Criteria for selecting an ASD process

The ultimate objective when selecting an ASD process is to obtain microbicidal efficacy against the strains used in the laboratory. The microbicidal efficacy of the device/product combination should be assessed under the conditions of use specified by the manufacturer. This efficacy with regard to the known targets handled in the laboratory can be measured through biological indicators selected as being representative,

Table 2. Example of biocidal products used for ASD

Product	Forms	Conditions of use	Advantages	Disadvantages
Formaldehyde	Liquid	3% to 10%	Broad spectrum of activity	Highly irritating, toxic, mutagenic, carcinogenic by inhalation
Formaldehyde	Gas	4 to 10 g/m3 18 - 22°C and 70% d'humidité	Broad spectrum of activity	
Glutaraldehyde	Liquid	2% Optimal pH: 8	Broad spectrum of activity	Irritant, toxic to the skin and respiratory tract. Activity greatly reduced in the presence of soiling.
Chlorine derivatives: Sodium hypochlorite, Sodium dichloroisocyanurate, Chloramine T	Liquid	Optimal pH: 6-7	Broad spectrum of activity	Aggressive. Toxic disinfection by-products. Activity reduced in the presence of soiling
Chlorine dioxide	Gas	Soluble in water	Broad spectrum of activity, Unlike hydrogen peroxide gas, it can tolerate a wide range of temperature and humidity	Produced *in situ*, corrosive
Peracetic acid	Liquid	Relatively unstable: decreases by 0.4% per month	Active at low concentrations in the presence of organic and inorganic soiling. Good candidate to destroy biofilms	Irritating to eyes and respiratory tract.
Hydrogen peroxide	Gas Liquid	Useable from 5 to about 35%. Relatively unstable.	In fumigation: faster and safer than formaldehyde. More stable than peracetic acid. Greater activity in the gas/liquid form	Depending on the procedure, may require humidity to be controlled at a low level. Some devices are expensive.

Focus

in terms of resistance of the micro-organisms handled by the laboratory. These biological indicators can then be used to validate the ASD process.

Criteria of the device/product combination

The correspondence between the device's power to diffuse the biocidal product and the volumes to be treated must be assessed. The upper and lower limits to the volume that can be treated by the device must be specified by the supplier. The microbicidal efficacy must be assessed for the device/product combination, under the conditions of use specified by the manufacturer. For this reason, it is important to obtain from the manufacturer full test reports that must be consistent with the claims and conditions of application. If the manufacturer makes any technical changes to the device (e.g. changes to the nozzle, pump, etc.), it must provide updated test reports. In terms of performance, the minimum log10 reductions expected by the manufacturers, according to the NF T 72-281 (2009) Standard, must be greater than or equal to, respectively:

- 5 log for bactericidal activity;
- 4 log for fungicidal activity;
- 3 log for sporicidal activity;
- for virucidal activity, if the falls in titres from the other European standards are used, this is set at a minimum of 4.

In addition to these levels of requirements specified by the standards, each laboratory should consider whether to adapt these performance levels to the risks associated with the biological agents handled.

Laboratory criteria

The room may be more or less complex in shape and have a varying degree of partitioning. Furthermore, the internal space may differ greatly, not only from one room to another but also in the same room, from one period to another and thus from one disinfection operation to another. Consequently, during the initial validation of the ASD application conditions described in the next section, and then during successive disinfection operations, a review of rooms and equipment is recommended in order to determine any changes and assess their potential impact. To promote air circulation for better diffusion of the disinfectant in the most inaccessible areas, aeraulic disinfection bypasses and/or fans carefully positioned in the TWRs may be used. Users must be especially vigilant to the risk of diffusion outside the room to be treated, given the toxicity of the biocidal product.

Procedure for validating an ASD process

ASD must be systematically validated (biological qualification) to ensure the efficacy of the process. It must undergo an initial validation before being used routinely.

It is important that operators and managers of the operation be qualified to apply and verify ASD, and to decide whether it is compliant. As each laboratory has its own organisation and configuration, it is difficult to see how this phase could be standardised. Only the main principles will be presented here.

Initial validation procedure

Initial validation involves ensuring that ASD is suitable for the room and the activities concerned, and that it meets the requirements set. It is a prerequisite to allowing a containment laboratory to become operational, and guarantees that, when used routinely, the process will be optimal. Therefore, this procedure must be validated, documented and formalised. All the critical parameters (temperature, humidity, time, etc.) should be verified and if possible recorded continuously throughout the ASD. Any changes to the layout of the room must be assessed with a view to possibly reconsidering the initial validation procedure. The initial validation phase includes a detailed plan of the containment laboratory specifying:

- the choice of biological indicators (BIs);
- the plan for positioning BIs;
- the plan for positioning chemical indicators, if necessary and if available;
- the plan for positioning ASD devices and fans, if necessary.

Choice of biological indicators

The choice of BIs is crucial because it is proof of the microbicidal efficacy of the process implemented. Among the biological criteria for selecting these indicators, the following should be taken into account:

- the nature of the micro-organisms to be tested that will be used as BIs;
- the nature of the medium on which these test micro-organisms are deposited;
- the possible presence of interfering material.

Regarding the choice of micro-organism, the microbicidal activity of the device/product combination should ideally be tested on the micro-organisms actually used in the laboratory. However, some micro-organisms cannot be tested because of their high pathogenicity and/or for technical reasons (this is the case with viruses, for instance). In this case, a literature review is needed to find representative micro-organism(s) from among non-pathogenic representatives of the same family, or for reasons of safety, highly resistant micro-organisms should be chosen whose resistance covers that of a very wide range of micro-organisms. For information, suppliers typically recommend using *Geobacillus stearothermophilus* and *Bacillus atropheus* as BIs. These two micro-organisms have long been used to test sterilisation methods using moist or dry heat. However, the use of commercial BIs occasionally has certain disadvantages:

- the preparation method for commercial BIs (drying on medium from spore suspensions placed in aqueous or alcoholic solutions) is not always representative of the target micro-organisms (such as viruses) that may be included in complex environments (culture media, excretions, biological fluids, faeces, etc.) likely to interfere with the disinfectants and reduce their efficacy;
- following the ASD operation, when the strips containing these BIs are returned to culture, they may contain residual disinfectant likely to inhibit germination and growth of vegetative forms.

The chosen solution could then be for the laboratory to manufacture its own biological indicators. Regarding virucidal activity, the limits inherent to some BIs should be taken into account (varying loss of viral titre on drying, difficulty producing a sufficiently large viral stock, no cell line for production and titration of virus, etc.). Selection of biological indicators, whether they are commercial or in-house, must also take into account the nature of the medium for the micro-organisms: the medium on which the BI is deposited should not interfere with the disinfectant. The medium used is normally made of stainless steel, or glass or plexiglass slides. It should not promote excessive adsorption of the disinfectant, which could inhibit the culture of the BIs. The last parameter to be taken into account in the choice of BIs, which would enable the efficacy of the chosen ASD process to be validated, is the presence of interfering materials. Indeed, in some cases, such as in facilities

Journal No. 12
Summer 2014

Focus

holding large animals or autopsy rooms, the microbicidal activity should be tested in "dirty" conditions. This is because there is interference between disinfectants with high chemical reactivity (oxidants, aldehydes, peracids, quaternary ammonium, etc.) and various organic media that can contain micro-organisms. For this reason, it is necessary to deposit the BIs on a strip containing an organic medium representative of the organic materials (secretions, excretions, faeces, blood, etc.) of interest to the operator. Usually this means using semi-skimmed milk at 1/20 dilution, 1% albumin or 1% yeast extract.

Once the choice of BI has been made, the ASD will be validated by calculating the reduction factor in the microbial population. This reduction is assessed by comparison with a BI that has not been exposed to the decontaminant. When using dehydrated BIs, it should be noted that during drying, 1 to 3 log losses of titres are not uncommon. The media may contain residual disinfectant that must be neutralised to avoid inhibition of micro-organism growth or a toxic effect on the cell system. In the latter case, gel filtration could be considered. Ideally, the process should be carried out three times to ensure reproducibility of results and that the results from the BIs meet the established compliance criteria. When the results are unsatisfactory, an analysis should be carried out to determine why the BIs were not completely inactivated. The following questions should be asked: is it related to a failure of performance by the process or the way it was applied? Is it related to a difference in the resistance of the BIs from one batch to another?

Plan for positioning biological indicators

This must be established taking into account the volume of the room, its shape, available space, MSCs, incubators and refrigerators, levels, specific critical areas, etc. Indicators should be positioned horizontally and vertically, in such a way that both sides of the strip can come into contact with the biocide. They should be placed in the parts of the laboratory most inaccessible to the product.

Plan for positioning chemical indicators

Where available, chemical indicators for detecting the presence of the decontaminating product should be used and positioned in the least accessible places in order to detect any possible heterogeneity in the decontaminating treatment.

Plan for positioning ASD devices and fans if necessary

This positioning plan should display the list of the devices used (name, serial numbers) to ensure that they are always positioned in the same place during routine ASD.

Routine checks

Each time ASD is implemented, it must undergo biological qualification. However, for routine use, it is no longer necessary to place the BIs in the most inaccessible places, they may be limited to critical areas where contamination is most common (handling areas, incubators, refrigerators, storage areas, areas of personnel traffic, etc.). This routine checking phase assumes that nothing has changed since the initial validation (no modification to the ASD process during upkeep or maintenance - no rearrangement of the room - no new micro-organisms handled in the laboratory, etc.). If this was not the case, the initial validation operation should be repeated. A fundamental point is verification of the compliance criteria for the disinfection cycle, which must be fully consistent with those used in the initial validation. This is why it is important that the temperature and humidity parameters be systematically recorded.

Focus

Practical guidelines for the implementation of the ASD procedure

This section offers a set of practical guidelines for the implementation of ASD. This list, based on the experience and discussions of the CMRBL, is not exhaustive.

Preparatory phase

During the preparatory phase, entering and leaving the BSL3 area shall be in accordance with the usual work procedures.

Operations	Description	People involved
Tidying and removing clutter	Put all equipment away in its place, if possible in cupboards or drawers. Remove paper and cardboard as well as small consumables not wrapped in double packaging. Empty water baths.	Personnel with BSL3 authorisation and trained in disinfection
Sealing doors	Seal exterior doors (interlocking security doors for equipment and personnel) with adhesive tape, if they do not have inflatable seals, to prevent micro-leakage.	
Protecting equipment that must not be decontaminated	Use hermetic packaging to protect equipment that must not be decontaminated and that will therefore not be accessible while the laboratory is open (e.g. microscope, computer equipment, etc.). Tape the doors of chambers containing biological material to prevent access. If possible, lock these chambers with keys or protection bars.	
Preparing materials and equipment	*To enable the decontaminant to come into contact with all areas, open cupboards, drawers, furniture under lab benches (if empty), incubators, centrifuges, etc.* *If it is compatible with the decontaminant, switch on the containment equipment (MSC, insulators, ventilated cabinets, ventilated rack) to ensure that the product passes through the filters.* *If the containment equipment is not compatible with the decontaminant, consider disinfecting the filters separately.* *Please note that switching on the containment equipment may also have an adverse affect on the product's diffusion in the room.*	
Surface cleaning	Thoroughly clean all surfaces (lab benches, worktops, MSC, incubators, etc.) using the detergent/disinfectant normally used for cleaning surfaces in the laboratory.	
Waste disposal	Take out all of the waste bins (after tidying, removing clutter and cleaning) and autoclave them. Remove liquid waste (water from incubators, water baths, etc.) either by autoclaving (liquid cycle) or *via* the effluent treatment plant by pouring them in the sink.	Personnel with BSL3 authorisation and trained in the use of incubators
Putting up signs	Display a sign on freezers, refrigerators and non-decontaminated equipment: "Do not open/use". Display a sign at the laboratory entrance explaining that entry is prohibited during the disinfection phase.	Personnel with BSL3 authorisation and responsible for application of the disinfection procedure
Positioning biological and chemical indicators	Prepare the indicators and position them according to the validated procedure. The number of BIs will depend on the volume of the room to be disinfected (usually between 3 and 5 for 10 m^2).	
Putting ASD devices in place	Check beforehand that the ASD devices are in working order. The ASD devices should be positioned according to the volume to be treated and the available space in the room. If possible, arrange them so they are visible from the outside in order to be able to check their operation. Check that the volume of decontaminant is adequate for the disinfection cycle.	
Putting fans in place	If the laboratory is not equipped with a disinfection "bypass", use fans to improve the diffusion of the disinfectant: note that they must be arranged as validated and described in the disinfection procedure.	
Sealing the entrance door	Leave the area to be disinfected and tape the entrance door giving access to the area.	

Disinfection phase

Once the operators are out of the room, the disinfection phase can begin. It is then necessary to switch off or bypass the ventilation system.

The ASD device will then operate for the time that was previously defined during the initial validation phase. When the contact time has elapsed, the ventilation system is restarted.

Focus

Post-disinfection phase

At this stage, when entering the laboratory, the operator must wear the normal PPE.

Operations	Description	People involved
Ventilation time	Generally 2 to 5 hours but depends on the product and process used. It is preferable to speak of air renewal rate. Each decontaminant manufacturer provides guidance in Volume/h.	Personnel with BSL3 authorisation and responsible for the disinfection procedure
Verifying that there is no remaining disinfectant	Measurement of the residual concentration of the disinfectant in the air by means of test tubes and/or a measuring device. Measurement of the concentration of disinfectant must give a result consistent with the requirements of the product's safety data sheet.	
Verifying the efficacy of disinfection	Recover the BIs and culture them; this operation is performed in the BSL3 laboratory. Remove the stoppered test tubes after surface disinfection and incubate. Check BI growth in the tubes daily for at least 5 days.	
Validating the disinfection	Disinfection can be validated if the BIs most exposed to the disinfectant remain negative after the required culture time and the chemical indicators are positive.	
Disinfection certificate	Produce a disinfection certificate that can be given to all companies working in the BSL3 for inspections and maintenance.	Laboratory manager + person in charge of the facility and of risk management + Laboratory director
Authorisation to reopen the laboratory	Remove the signs prohibiting entry to the containment area.	Personnel with BSL3 and disinfection procedure authorisation

Protection and safety of personnel

Disinfection preparation phase	Description
Type of protection	Protection against biological risks: PPE normally used in the BSL3 laboratory.
Safety instructions	Put up the "disinfection in progress" signs. Follow the instructions relating to lone workers: a LWP system must be worn or mandatory presence of 2 people inside and one person outside who can intervene in an emergency.
Disinfection initiation phase	**Description**
Type of protection	To start the device: normal PPE but with a protective mask against chemical contamination fitted with a chemical filter suitable for the disinfectant used (e.g. single-cartridge gas mask with panoramic visor and chemical filter).
Safety instructions	Work in pairs monitored from the outside by at least one person who can intervene if necessary.
Disinfection verification phase	**Description**
Type of protection	PPE suitable for the biological risk and chemical mask suitable for the disinfectant used (see above).
Safety instructions	Work in pairs monitored from the outside by at least one person who can intervene if necessary.

Conclusion

Disinfection of surfaces in a containment laboratory is a complicated yet important operation, both for the safety of users and for the laboratory environment. It must be carried out meticulously and methodically, with the aim of achieving zero contamination risk. Given the many parameters involved in this operation, we have seen that there is no universal process or method. Only a very good knowledge of the situation of the laboratory concerned and its operation, along with the training of qualified personnel, will guarantee the success of the operation. Carrying out an ASD operation is relatively onerous and expensive. Therefore, once it has been validated, we recommend writing down the entire protocol in detail to ensure that the operation can easily be reproduced. Document traceability is important because it is often an aid to reconstruction of events in case of an incident.

Focus

Definitions and acronyms

A3 animal housing: Biocontainment level 3 animal housing in accordance with the Decree of 16/07/2007

Aerosol: Suspension of solid or liquid particles in a gas with a negligible settling rate

Afnor: French Standards Institute

AHU: Air handling unit

ANSES: French Agency for Food, Environmental and Occupational Health & Safety

ANSM: French National Agency for Medicines and Health Products Safety

ASD - Airborne Surface Disinfection: operation with a temporary result that reduces contamination of inert environments or surfaces by micro-organisms to an acceptable level, depending on the objectives set by the risk assessment. This operation, performed by means of a process whose active ingredients are chemical agents in gaseous or mist form, is intended for disinfecting surfaces in a given volume, regardless of their orientation.

Automatic disinfection process: Process that diffuses a substance in gaseous or mist form, solid or liquid, from an emitting source, without human presence.

BI: Biological indicator

Biocide: Preparations containing one or more biocidal active substances presented in the form in which they are delivered to the user, and intended to destroy, deter, or render pests harmless, to prevent their action or to combat them in any other way, by chemical or biological action. A disinfectant used for ASD is a biocide.

Bio-cleaning: Operation that combines cleaning and disinfection

Biological agents: Micro-organisms, including those obtained by genetic engineering, cell culture and endoparasites, whether pathogenic or not.

Biological pathogens: Biological agents capable of causing infection, allergy or toxicity or otherwise constituting a risk to human health.

BSL3: Biological safety level for handling group 3 micro-organisms as defined by the Decree of 18 July 1994.

BSL3 laboratory: Biosafety level 3 containment laboratory

Chemical disinfection: according to the EN 15889-1 Standard: "action of one or more chemicals whose main aim is to be a microbicide".

CMR: Carcinogenic, mutagenic, reprotoxic

CMRBL: Comittee for the Control of Laboratory Biorisks

Contact time: Time required to reach the expected efficacy.

Containment: Series of technical measures and actions aimed at keeping a biological agent or other entity within a given space.

Containment area: Area built and used (and equipped with a suitable air treatment and filtration system) to prevent the external environment being contaminated by biological agents from this area.

Contaminant: Any particulate, molecular, non-particulate or biological entity likely to produce an adverse effect on a product, process, organism, or on the environment in general.

Contamination: Phenomenon of interaction by contact between two entities, one being the contaminant and the other the target, entailing disturbance of the target and whose consequences can be diluted over time.

CSP: French Public Health Code

Disinfection: Operation consisting in reducing the number of micro-organisms in or on an inert matrix, achieved by the irreversible action of a chemical or physical process on their structure or metabolism, at a level deemed acceptable for a defined objective.

According to the AFNOR NF T 72-101 Standard: disinfection is "an operation with a temporary result that can eliminate or kill micro-organisms and/or inactivate viruses carried by contaminated inert matrices, depending on the objectives set".

Disinfection bypass: Aeraulic system enabling closed-circuit laboratory ventilation

Dispersion: Dissemination of micro-droplets in the air.

Dispersion time: Time required to reach a defined concentration of the product in a given volume.

IARC: International Agency for Research on Cancer

ID_{50}: Dose infecting 50% of target tissues or species.

Inactivation: Partial or complete destruction of a given activity or destruction of a microbiological system.

Infected: Contaminated by foreign biological agents that can multiply in a matrix and may or may not reproduce there.

LWP: Lone worker protection

MA: Marketing authorisation

MEL: Maximum exposure limit

Micro-organism: Any microbiological entity, whether cellular or non-cellular, capable of reproducing or transferring genetic material

MSC - Microbiological safety cabinet: ventilated enclosure designed to protect the user and the environment from hazards related to aerosols when handling potentially hazardous and hazardous micro-organisms, with the air filtered before being released into the atmosphere.

Non-compliance: Non-fulfilment of a requirement

Procedure: Description of the operations to be carried out and precautions to be taken in an area, directly or indirectly related to the micro-organisms or toxins.

PPE: Personal protective equipment

Qualification of equipment, facilities, a room: Operation seeking to demonstrate that the equipment/facilities/room function properly and actually give the expected results.

PT2 - Product Type 2: Private and public health area disinfectants and other biocidal products

REACH: European Regulation on the registration, evaluation, authorisation and restriction of chemicals. It entered into force on 1 June 2007.

Risk: Probability of occurrence of a hazard causing harm and the degree of severity of this harm.

TWR - Technical Work Room: rooms in which samples, bodies and animals – which have been or are likely to be contaminated with biological pathogens – are handled, as well as rooms in which biological pathogens are intentionally handled.

Validation: Establishment of proof, in accordance with the principles of good manufacturing practices, that the implementation or use of any process, procedure, equipment, raw materials, packaging article or product, activity or system can actually achieve the desired results.

EuroReference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Research

AMCRA : center of expertise on antimicrobial consumption and resistance in animals in Belgium

Bénédicte Callens (scientific collaborator, AMCRA - benedicte.callens@amcra.be) and Evelyne De Graef (coordinator, AMCRA)

Keywords: network, antimicobial resistance, animal, Belgium

Objectives and achievements

AMCRA is a Belgian initiative that was launched in January 2012. From the 'One World, One Health' perspective, AMCRA aims to fight against the selection and spread of antimicrobial-resistant bacteria by promoting reduced and responsible use of antimicrobial agents in veterinary medicine in Belgium. This requires joint action through an industry-wide approach, together with the authorities and all relevant stakeholders: veterinary practitioners, agricultural organizations, the pharmaceutical industry, compound feed manufacturers and the scientific world. In order to accomplish this mission, several strategic objectives were achieved in 2012 and 2013, and more objectives are planned for 2014.

A first and crucial objective achieved involved drafting a study report on the establishment of an electronic database on antimicrobial consumption per animal sector and category, at the veterinarian and farmer levels. The quantification of antimicrobial use in veterinary medicine and corresponding identification of high-level users is a major priority. Indeed, in January 2014 a Belgian swine-farming quality system initiated a data collection programme for providers of antibiotics in the swine-farming industry, which covers approximately 60% of swine farms in Belgium. Further data collection systems are planned for the poultry and veal calf industries, two other 'high use' sectors. With regard to data collection, methodology and calculations, special attention should be paid to harmonization between animal sectors and countries in order to keep data comparable. In addition, data analysis output must produce reliable 'herd-orientated' results as they are the basis for herd-specific recommendations.

A second objective for 2012 was to study the possibility of using zinc oxide (ZnO) at pharmacological doses in swine feed as a possible alternative to use of antimicrobial agents. The report on ZnO included an assessment of environmental issues and questions concerning dosage. Several temporary registrations of pharmaceutical doses of ZnO in feed were granted in Belgium between August and September 2013. In this respect, Belgium is the 11th country in the EU with a (temporary) approval. Also, an agreement has been reached between the Belgian government, compound feed manufacturers and feeding companies to reduce the amount of zinc used as a feed additive from 150 ppm to 110 ppm in the fattening phase. Even though the dosing schedule during fattening leads to a total decrease of 4.5% of the environmental Zn burden, the search for alternatives to both antibiotics and heavy metals should be continued, as resistance to Zn might occur in bacteria and Zn is a potential selector for methicillin-resistant *Staphylococcus aureus* (MRSA).

Furthermore, in 2012 and 2013, several working committees with relevant representatives of all stakeholders concerned were set up per animal species/group (swine, poultry, veal calves, beef cattle and dairy cattle). These committees made proposals for self-regulation of the animal sectors with respect to the use of antimicrobial agents in veterinary medicine. These proposals include measures concerning production and trade, aimed at veterinarians and farmers. The working committees also worked closely on the establishment of national guidelines on the careful and rational use of antimicrobial agents in livestock animals. The multidisciplinary expertise of all stakeholders involved resulted in the establishment of well-structured and valuable practical guidelines with both general information regarding good herd health, and the correct use of antimicrobial medicines as species-specific information with specific therapeutic recommendations depending on the disorder ('formularies'). Currently, guidelines for pets (dogs and cats) and for horse practitioners are being developed.

Moreover, other challenging aspects were dealt with, for example storage of medicines by veterinarians and farmers which led to drafting of an opinion on the need to hold a stock of antibacterial agents, etc., and the conditions for doing so.

More activities are planned, such as drafting proposals to broaden legislation on the prescription and use of medicated feed with a view to centralised electronic data collection on the prescription and use of medicated feed. Also, guidance documents are planned concerning the implementation of alternatives that could reduce the use of antibiotics, with special attention to vaccines (e.g. the correct application of vaccination programmes) and biosecurity measures, for all animal species/groups.

Other challenging objectives are on the horizon, such as discussions with the authorities and the different stakeholders about targets for a specific percentage reduction in antibiotic use in livestock animals by a certain year, or concerning restrictions on the use of the critically important antimicrobials for human medicine in food-producing animals. Furthermore, it is and will remain essential to continue encouraging the different animal sectors and quality labels to provide specifications concerning the implementation of self-regulating measures, and the use of animal health guidelines and formularies.

Many actions have been implemented these last 2 years, but there is still a long way to go. It is only through continuous and committed participation of all stakeholders that the challenging goals of this industry-wide and unique initiative can be met. In this respect, a central aspect of AMCRA's mission is to continue informing veterinarians, farmers and animal owners on rational use of antibiotics by developing information and awareness campaigns and giving presentations for veterinarians and farmers locally.

Would you like to stay informed about our progress and activities? Please visit our website at www.amcra.be. You will find all our guidance documents and reports on finalised objectives, as well as guidelines for veterinarians and farmers, and upcoming events in regard to awareness campaigns.

Euro Reference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Research

ACTEOLab-*Salmonella* : more than a database for the French *Salmonella* network, a tool for monitoring *salmonellae* of non-human origin.

Renaud Lailler[1] (renaud.lailler@ANSES.fr), Isabelle Berta-Vanrullen[2] (isabelle.berta-vanrullen@ANSES.fr), Louis-Ziad Alexandre[3] (Louis-Ziad.Alexandre@ANSES.fr)
1. ANSES, Laboratoire de sécurité des aliments, Maisons-Alfort, France.
2. ANSES, Direction scientifique des laboratoires, Maisons-Alfort, France.
3. ANSES, Direction technique et informatique, Maisons-Alfort, France

Since 1997, the ANSES Laboratory for Food Safety has been coordinating and leading the French network for the epidemiological surveillance of *Salmonella* strains of non-human origin. This network currently comprises some 140 laboratories throughout France.
Work has begun on modernising the network's information system to extend the database's functional capabilities and storage capacities while tailoring it to EFSA standards (standard sample description available at www.efsa.europa.eu/fr). ACTEOLab-*Salmonella* is a web application used to centralise and transfer data. Geared to the *Salmonella network's* operational epidemiology needs, it is designed to apply epidemiology tools to enhance the relevance and responsiveness of the surveillance system on *Salmonella* strains of non-human origin throughout France.

Keywords: data bank, non-human *Salmonella*, network

ANSES, the French Agency for Food, Environmental and Occupational Health & Safety, was entrusted with creating or contributing to the creation of scientific and technical databases within its fields of expertise by Ministerial Order No. 2010-719 of 28 June 2010, issued by the French Ministry of Food, Agriculture and Fisheries.
The Laboratory focuses on biological and chemical contaminants that may affect consumer health. It participates in all ANSES's missions within the area of food safety, from reference missions to research, vigilance, epidemiology or scientific and technical expert assessments. The laboratory therefore uses serological and molecular methods to characterise the bacteria responsible for collective or sporadic outbreaks of foodborne illness. This activity is crucial to investigating food contamination and tracking down its origin. The laboratory also monitors and detects emerging hazards such as resistance to antimicrobials or emerging bacterial clones. It holds several national and European reference mandates concerning bacteria identified as major foodborne human pathogens, including *Salmonella*.
In conjunction with the ANSES Ploufragan-Plouzané National Reference Laboratory (NRL) for *Salmonella*, the LSAL conducts and coordinates characterisation activities for *Salmonella* strains of non-human origin. It coordinates the *Salmonella* network, a group of 140 public or private laboratories performing food and veterinarian analyses. These laboratories collect strains isolated in a variety of contexts, including in-house checks carried out by the food-processing industry, official monitoring and inspection plans, investigations or food alerts. At the same time, they collect epidemiological information relating to these isolates. The network reveals the diversity and spatiotemporal evolution of isolated serovars throughout the food chain (Lailler, 2012). It is a source of information on rare serovars or serovars not covered by regulations, and may alert health authorities when necessary (Danan, 2012).

ACTEOLab-*Salmonella*, the *Salmonella network database*

The microbiology analysis data produced within the network do not only provide scientific and technical support to partner laboratories. Their full value is revealed when processed for epidemiological surveillance purposes. ACTEOLab-*Salmonella* originated from the need to update the information system used since 2001 to cope with increasing data volumes, coordinate the *Salmonella* network and produce health indicators for surveillance purposes. The project was initiated in 2012 and work continued throughout 2013.
ACTEOLab-*Salmonella* is an application designed to centralise and transfer data for operational epidemiological surveillance purposes. Run by the *Salmonella* network (www.ANSESpro.fr/reseausalmonella), it helps guide laboratory missions under the national salmonellae alert and surveillance system.
Once data have been centralised and processed, health indicators of particular use to risk managers may be generated. These could include trends in isolation or the emergence of strains critical to human health. ACTEOLab-*Salmonella* is therefore a key component in the French food-chain surveillance system. It complements the monitoring data on food from other sectors, such as data on chemical contamination from total diet studies.
Through this web application, the Laboratory aims to (i) develop, manage and maintain an information system compatible with ANSES requirements, (ii) remain in close contact with other laboratories in the network and its partners, (iii) combine the results from different teams of biologists typing strains, and (iv) communicate with users and partners.
Each laboratory keys into ACTEOLab-*Salmonella* the results of analyses and epidemiological data received in the post or produced within the laboratory. In the long term, ACTEOLab-*Salmonella* should be able to directly integrate data from network partners in electronic format. The application can be used to monitor the processing of analyses and perform multiple-criteria searches. It complies with the EFSA data standardisation classification (standard sample description, v.2) (EFSA, 2013).

Development of ACTEOLab-*Salmonella*

Initiated in 2012, the ACTEOLab-*Salmonella* project required scheduling various development phases. Each one was

Research

characterised by detailed specifications based on the definition of specific needs, followed by a technical and functional acceptance test stage to check proficiency. Following implementation of each phase, there was a period for maintenance and corrective updates.
The 'Agile' development methodology applied to this project is an incremental method iterated over short periods. It led to a reduction in development costs, satisfied profession-oriented needs, constituted a library of reusable modules for future web applications and, significantly, removed the risk that detailed specifications phases represent in fixed-price methods. In-house human resources guided a service provider in producing a prototype application which revealed actual needs and enabled a shift in focus during the specifications phase.

Finally, ACTEOLab-*Salmonella* meets several objectives:

• Securing network data

The first ACTEOLab-*Salmonella* development step was completed in July 2013. All the data collected are currently hosted in a shared Oracle database offering the benefits of related technical environments at no extra cost. Older data currently centralised in two MS Access databases are being migrated to ensure their long-term accessibility.

• Modernising the web tool and interface

The ACTEOLab-*Salmonella* application was developed with full web technologies using free environments offering dynamic graphic interfaces. The migration of data stored in the previous database required transposing the existing reference system into EFSA's standard thesauruses. Analysis reports are processed through an ANSES cross-disciplinary document management tool.

• Rendering the tool generic

ACTEOLab-*Salmonella* was designed with generic service-oriented programming technologies for greater reusability. The project is therefore a foundation upon which future epidemiological surveillance projects may be built.

• Data processing for epidemiological surveillance needs

This future development stage is a natural corollary leading to the ultimate goal and value-added exploitation of the microbiology analysis data produced. Although at its current development stage ACTEOLab-*Salmonella* cannot yet offer surveillance tools, comparable surveillance information systems such as RESAPATH include features such as the automatic generation of health or performance indicators presented in the form of a dashboard which is both easy to use and interpret (Enki, 2013; Sorbe, 2011; Hulth, 2010; Weisent, 2010). This dashboard could be specifically geared to the challenges facing each particular user of the monitoring system.

A tool to coordinate and direct the laboratory network

It is accepted that the key contributors driving the performance of a surveillance system are the coordination and regular assessment of network operation, harmonisation of the analytical methods and standards used for exchanging data, and the communication tools and resources used (Dufour & Hendrikx, 2011; Lailler, 2012).
Developed information systems obviously rise to such challenges. They help by providing a standardised approach to diagnostic values so that plug-in algorithms can check data quality and generate health and performance indicators. These IT systems are also used to key in, store and transmit information by exchanging computer data. Through the web interface and solutions ensuring data confidentiality, ACTEOLab-*Salmonella* can also provide feedback to the multiple partners of the surveillance network. The ease with which information can be sent to network partners, in the form of dashboards for example, or the support provided for annual inventories of characterised *Salmonella* serovars are crucial for motivating players and boosting the performance of the whole epidemiological surveillance network.

A tool for health and epidemiological surveillance rooted in a recognised network

According to the 'one health' approach described by Bousfield & Brown (2011), the health monitoring of salmonellae in France requires real-time access to an instant picture of circulating serovars and appropriate health indicators in the human population and the food supply chain, from the farm to the fork. It is also vital that this picture be monitored over time. Such an approach is needed to protect public health in France by quickly detecting an emerging pathogen so as to restrict its dissemination in the population, for instance. Finally, it is vital to collect information on serovars and the accuracy of the epidemiological values computed to study the risk factors associated with salmonellae.
However, the application of methodological actions needed for surveillance is currently inadequate. Repeated observations point to limits in surveillance protocols and insufficient sampling that must be improved to obtain more representative and robust data. There is also room for improvement in the processing of data produced by analysis and the computerised data management tools available. Finally, another finding was poor coordination in the area of epidemiological surveillance.
Centralising different kinds of data within the same IT system will enable newly emerging pathogens to be identified quickly and will provide support to investigations into outbreaks.
Even outside emergency situations, ANSES regularly produces synopses describing the data available in the *Salmonella* network database. Epidemiological data are collected for each strain and complemented by the results of analyses obtained once the different characterisation methods have been applied. The goal is to specify the context and origin of the microorganisms isolated and analysed. These data are very valuable for epidemiological surveillance and help risk managers facing an emergency or health alert. The data produced and collected by the ANSES LSAL relate to strains of non-human origin and originate from different stages of the food chain. They therefore round out the information available from the National Reference Centre and the French Institute for Public Health Surveillance.

Looking ahead ...

ACTEOLab-*Salmonella* is currently being used in the LSAL by the team responsible for coordinating the *Salmonella* network. Its roll-out in summer 2013 secured the data collected by the network since 2001.
The application should shortly be opened up to network partners and integrate new modules and interfaces dedicated to the epidemiological surveillance of salmonellae of non-human origin.
The IT development strategy of ACTEOLab-*Salmonella* is such that it will be able to evolve in keeping with technological developments in the characterisation of foodborne microbial contaminants, such as whole genome sequencing. More

Research

generally, the application's generic nature opens up prospects for its adaptation and use for other pathogens, other surveillance networks and other organisations.

Acknowledgements

We would like to thank the *Salmonella* network coordination team—A. Brisabois, V. Leclerc, F. Moury, V. Morel, C. Oudart and C. Piquet—for their daily commitment to *Salmonella* reference activities and their active participation in the development and implementation of ACTEOLab-*Salmonella* in first the CEB characterisation and bacterial epidemiology unit then the SEL unit of the Laboratory.
Our thanks also go to the attentive and available IT team deeply involved in this project (E. Frigot, JC. Gnanou and MC. Poupard). We would also like to thank the ANSES administrative and financial team for their support, and Pascal Hendrikx for reviewing this article.

References

Bousfield B. & Brown R. 2011. One world one health. *Veterinary Bulletin - Agriculture, Fisheries and Conservation Department Newsletter* No. 1 Issue No. 7.

Danan C., Agbessi A., Cabassut G., Moury F., Guyot M., Talleu L., Salah S., Chemaly M. 2012. Surveillance des salmonelles isolées de la chaîne alimentaire en France. B.E.H. (Special issue, 9 May 2012, Microbiological hazards in food products of animal origin: monitoring and evaluation) p. 30-33.

Dufour B. & Hendrikx H. 2011. Surveillance épidémiologique en santé animale. Editions QUAE. Paris.

EFSA. 2013. Standard Sample Description ver. 2.0. *EFSA Journal*, 11:3424, 114 pp., doi:10.2903/j.efsa.2013.3424.

Enki D.G., Noufaily A., Garthwaite P.H., Andrews N.J., Charlett A., Lane C., Farrington C.P. 2013. Automated Biosurveillance Data from England and Wales, 1991-2011. *Emerg Infect Dis.* 19:35-42. http://dx.doi.org/10.3201/eid1901.120493

Sorbe A., Chazel M., Gay E., Haenni M., Madec J.-Y., Hendrikx P. 2011. A simplified method of performance indicators development for epidemiological surveillance networks – Application to the RESAPATH surveillance network. *Rev Epid Santé* Pub, 59:149-158.

Hulth A., Andrews N., Ethelberg S., J Dreesman, D Faensen, W van Pelt, J Schnitzler. 2010. Practical usage of computer-supported outbreak detection in five European countries. *Eurosurveillance*, 15:1-6.

Lailler R., Moury F., Granier S.A., Brisabois A. (2012). The *Salmonella* Network, a tool for monitoring *Salmonella* "from farm to fork". *Euroreference, n°8*, winter 2012.

http://www.ANSESpro.fr/euroreference/Documents/ER08-Reseaux-*Salmonella*EN.pdf

Weisent J., Seaver W., Odoi A., Rohrbach B. (2010) Comparison of three time-series models for predicting campylobacteriosis risk. *Epidemiol Infect*. 138, 898–906.

EuroReference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Research

Purpose and overview of results of the Vigimyc Network for the epidemiological surveillance of mycoplasmoses in ruminants in France

F. Poumarat[1,2] (françois.poumarat@ANSES.fr), N. Jarrige[3], F. Tardy[1,2]
1. ANSES, Lyon Laboratory, Joint Research Unit on mycoplasmoses in Ruminants, Lyon, France.
2. Université de Lyon, VetAgro Sup, Joint Research Unit on Mycoplasmoses in Ruminants, Marcy-l'Etoile, France.
3. ANSES, Lyon laboratory, Epidemiology Unit, Lyon, France.

Mycoplasmas cause several diseases in ruminants including three that are registered as diseases of concern internationally by the World Organisation for Animal Health (OIE). The Vigimyc Network was established in France to monitor the status of these regulated diseases and of other economically harmful mycoplasmoses. Vigimyc maintains a collection of strains that is representative of the epidemiological picture nationally and this collection is regularly used to develop and validate diagnostic tests and for various studies aimed at improving knowledge of mycoplasmas, including their antimicrobial susceptibility and their pathogenicity.

Introduction

The class Mollicutes that gathers bacteria characterised by their small size and absence of a cell wall, is essentially represented in animals by species of the *Mycoplasma* genus. In ruminants, about forty species and sub-species of mycoplasmas have been described, including some that are pathogenic.

Three mycoplasmoses have economic consequences and an impact on trade that is sufficiently serious to warrant international control measures and classification of these diseases by the World Organisation for Animal Health (OIE). On the one hand, there are two «exotic» mycoplasmoses that pose a threat of re-emergence or emergence in France: contagious bovine pleuropneumonia (CBPP) and contagious caprine pleuropneumonia (CCPP) and, on the other hand, contagious agalactia (CA) that is widespread in southern Europe. CBPP is caused by *Mycoplasma (M.) mycoides* subsp. *mycoides* (*Mmm*). The disease originated in Europe and became a major worldwide panzootic in the 19th century. It is still highly prevalent in Africa and sporadic in Asia. In Europe, the implementation of an eradication programme in the 1980s and 1990s following widespread resurgence has led to no outbreaks having been recorded since 1999. CCPP is caused by *M. capricolum* subsp. *capripneumoniae* (Mccp). It was thought that this disease was limited to north-east Africa, but improved diagnostic methods have demonstrated that it is far more widely spread through Africa and Asia up to the borders of Europe, and that wildlife is possibly infected, particularly zoo animals. CA is a complex syndrome characterised by mastitis, arthritis, pneumonia and septicaemia. It is present worldwide with a strong impact in the Mediterranean area. It can be caused by several mycoplasmas: *M. mycoides subsp. capri (Mmc)*, *M. capricolum* subsp. *capricolum (Mcc)*, *M. putrefaciens* in goats and *M. agalactiae* in sheep and goats.

Although they are not listed by the OIE, *M. bovis* mycoplasmoses have become significant with the development of trade and herd mixing related to modern cattle farming. They manifest in a highly proteiform manner with mastitis, arthritis, otitis and pneumopathies. Bronchopneumonia in young cattle poses problems internationally and mastitis or even otitis are becoming frequent and an economic burden in some countries, particularly in North America.

In addition to these major diseases, other mycoplasmoses are beginning to cause concern. *M. ovipneumoniae* in small ruminants is considered to be an important factor in respiratory disease in some countries. *M. leachii*, reported sporadically in Europe, has been found to be highly pathogenic in cattle in China and in Australia (arthritis, abortions, mastitis). *M. canis* and *M. alkalescens* are also thought to be pathogenic in cattle. Moreover, many other saprophytic mycoplasma species are found in ruminants and are sometimes abundant. Isolation of a mycoplasma is therefore of no clinical relevance unless the specific species or sub-species is identified.

In order to monitor these diseases, we created an epidemiological surveillance network for mycoplasmoses in ruminants in France in 2003, the Vigimyc Network. This article describes the organisation of this network, presents an overview of results obtained over the last 5 years, and the scientific benefits resulting from Vigimyc.

Objectives and organisation of the Vigimyc Network

Vigimyc was originally designed as a diagnostic support service to stimulate diagnosis of mycoplasmoses, but has subsequently evolved and now covers a significant part of the country, and can therefore be considered a surveillance network, despite certain methodological limitations.

Vigimyc has the following objectives:

- to identify mycoplasma species isolated in ruminants;
- to determine the epidemiological picture and follow-up on mycoplasmoses in ruminants across France, particularly those that are listed by the OIE;
- to detect any emergence of new mycoplasma species or variants;
- to share scientific and technical data regarding mycoplasmas;
- to build up and make use of a collection of nationally representative strains.

Vigimyc is administered by ANSES-Lyon Laboratory and supervised by a steering committee made up of representatives from all the stakeholders in the network: participating laboratories, public authorities, practicing veterinarians, farmers and scientists.

Vigimyc is a «passive» surveillance network since the decision to test for mycoplasmas is solely the initiative of the practicing veterinarian. Mycoplasma detection by subculture from clinical specimens is carried out by departmental veterinary diagnostic

Research

laboratories with some technical guidance from the network which is regularly updated during inter-laboratory testing. When mycoplasmas are isolated, the primary culture is sent to us for identification, along with a standard registration sheet containing background data on the sample. Identification is carried out using dot immunobinding on membrane filtration (Poumarat *et al.,* 1991). Each primary culture is tested using a battery of hyperimmune sera representative of the main mycoplasma species found in ruminants and a specific monoclonal antibody against the agent causing CBPP. If the result is ambiguous or negative, additional testing is carried out by specific PCR protocols or PCR and/or sequencing of various "housekeeping genes". The result is then forwarded to the requester laboratory. All epidemiological data and identification results are centralised in a database and a summary is provided annually to the participants and members of the steering committee.

Current trends concerning mycoplasmoses in ruminants observed through Vigimyc over the last five years

Overall analysis

Key figures are presented in Tables 1 and 2. Over the period 2009-2013, 46 veterinary diagnostic laboratories participated in the network. A total of 1938 primary mycoplasma cultures from 1526 outbreaks were sent to ANSES Lyon, resulting in identification of 2105 isolates, taking into account mixes of species. These isolates were from 77 different French *départements*, i.e. 80% of the country, and 44% involved cattle, 38% goats, 12% sheep and 6% wildlife, primarily ibex.

In cattle, most isolates were from respiratory diseases in young animals and M. bovis was the most commonly identified mycoplasma. In goats, the isolates were mainly from contagious agalactia cases, with *Mmc, Mcc* and *M. putrefaciens* the most commonly isolated mycoplasmas. In sheep, isolates came primarily from respiratory disorders in lambs, with increasing isolation of *M. ovipneumoniae*. In mountain ungulates, isolates were either from pneumopathy lesions with an *M. agalactiae* characterisation, or from nasal or ear swabs, indicating high carrier levels of *M. feriruminatoris* in healthy animals.

Two non-pathogenic mycoplasma species, *M. bovirhinis* in cattle and *M. arginini* in all ruminants, were frequently isolated, alone or in combination, and are of no diagnostic significance.

There are no major changes since the last overview for 2003-2008 (Chazel *et al.,* 2010), with the exception of increased isolation of *M. ovipneumoniae* in sheep.

Picture for mycoplasmoses listed by the OIE

The specific agent for contagious bovine pleuropneumonia (CBPP), which was tested for systematically, was found in no animal species, whether cattle, its usual hosts, or small ruminants that can be occasional hosts.

There are three types of contagious agalactia (CA) caused by M. *agalactiae* in France: one in sheep, limited to the milk-producing area in the western Pyrenees and increasing sharply since 2006; one in goats that occurs sporadically across the country (11 outbreaks in eight départements between 2009-2013); and one in wildlife found in the Alps following an episode of mortality related to bronchopneumonia in ibex and chamois populations.

Caprine CA caused by *Mmc, Mcc* or *M. putrefaciens* has been found to be highly prevalent. It is mostly caused by *Mmc*, but the annual rate of *Mcc* isolation fluctuates significantly (higher than Mmc in 2013). *M. putrefaciens* is less frequent and is mainly related to mastitis. *Mmc* is also found sporadically in cattle and sheep. Some severe clinical forms of CA are very similar to contagious caprine pleuropneumonia (CCPP). Because it is difficult to grow *Mccp* on commercially available media, these outbreaks could go undetected. As a result, corresponding information has been distributed widely via Vigimyc so that any outbreak of serious pneumopathy along with high morbidity and mortality in goats would be reported to ANSES. The Agency would then be able to carry out a specific PCR test for CCPP directly on pleural fluid or the lung tissue without prior enrichment. In this way, two suspected cases were registered between 2009 and 2013 but were found to be related to *Mmc*.

Picture for *M. bovis* mycoplasmoses

M. bovis is the most commonly isolated mycoplasma in cattle in France but it is mainly found during pneumopathies, with an overall prevalence estimated at 15% on the basis of a one-off survey in 2013 among Vigimyc laboratories. The other clinical forms, mastitis, arthritis and otitis are far more infrequent. Only four outbreaks of mastitis across four départements were reported between 2009 and 2013. This very low to non-existent incidence does not appear to be an under-estimation bias since a study involving systematic testing on bulk tank milk in the Rhône-Alpes region arrived at the same conclusion (Arcangioli *et al.,* 2011). Arthritis is often associated with respiratory diseases while sporadic otitis outbreaks are starting to be identified by Vigimyc (eight outbreaks in three *départements* between 2009 and 2013).

Other mycoplasmoses

Until recently, *M. ovipneumoniae* was rarely isolated, even though very frequent co-infection with *M. arginini,* a fast-growing mycoplasma, could mask this infection. However, since 2010 and even though procedures have not changed, a much higher number of cases has been found in respiratory disease in small ruminants. Two hypotheses have been put forward to explain this progression: either evolution of the strains of interest, or a change in the type of samples being studied. Samples currently come primarily from lambs grouped for finishing, with the concentration of animals promoting high infection pressure.

Two species, *M. canis* and *M. alkalescens*, have emerged in cattle and are developing strongly in the United Kingdom and in some other European countries. They are thought to be involved in respiratory disease and for *M. alkalescens*, also in arthritis and mastitis. A retrospective study on older collections at ANSES has shown that these mycoplasmas have been present in France for a long time, with the oldest isolations dating from 1965 and 1993. Moreover, no real progression has been found in France for a decade.

No isolate with a profile indicative of *M. leachii* was detected in cattle, sheep or goats.

Scientific use of biological material generated by Vigimyc

Although the collection of mycoplasma strains assembled as part of Vigimyc cannot claim to be a biological resource centre, it does constitute a notable representative selection of the epidemiological picture in France over time, and reflects the biological diversity of mycoplasmas in ruminants. It enables method development and validation not only for diagnosis, surveillance and molecular epidemiology, but also up-stream, for analysis of changes and virulence factors in mycoplasmas.

Research

Development and validation of detection and identification methods for mycoplasmas

Since strains are constantly evolving and diagnostic techniques constantly improving, stakeholders working within the frame of Vigimyc need to regularly validate detection and identification methods for mycoplasmas. Indeed, degrees of specificity and versatility of techniques are sometimes called into question when they are assessed using large numbers of diverse samples, such as those provided by Vigimyc (Le Grand *et al.,* 2004; Marenda *et al.,* 2005). These difficulties are related to the high level of genomic plasticity in mycoplasmas (Marenda, 2014) which leads to sometimes significant diversity within (sub) species, despite sometimes close phylogenetic relationships between (sub)species. For example, denaturing gradient gel electrophoresis (DGGE) of sequences of 16S rRNA previously amplified by PCR, a technique that is widely used in the UK as the new universal method to identify mycoplasmas, has been found to be insufficiently selective for certain (sub)species that are very closely related phylogenetically in the more complex French epidemiological context (Tardy *et al.,* 2008). In contrast, MALDI-TOF-type mass spectrometry, tested more recently, appears to be very promising (Pereyre *et al.,* 2013). This technique can identify microorganisms by comparing their predominant protein profiles with an array of reference spectra. The Vigimyc collection is an excellent tool for regularly verifying the completeness and accuracy of this array.

Regular efforts are also made to develop new diagnostic techniques that are suited to the national epidemiological picture. For instance, a real-time PCR method, able to detect and identify the four etiological agents of CA simultaneously, has been developed and made available as a commercial kit as part of a partnership with a private company (Becker *et al.,* 2012). Furthermore, a PCR technique used for health watch has been designed to unequivocally distinguish between *Mcc*, frequently isolated in caprine CA, *M. leachii,* and above all *Mccp*, the agent underlying CCPP, through direct detection in clinical samples (Maigre *et al.,* 2008).

Molecular subtyping of strains

Molecular subtyping can be very useful in epidemiology and in disease control. The collection of strains stemming from Vigimyc is very useful in this way to compare strains from different times, hosts, diseases and regions. For example, subtyping the various isolates of *M. agalactiae* from our collection, specifically by Multiple Locus Variable number tandem repeat Analysis (MLVA) and macro-restriction followed by Pulsed Field Gel Electrophoresis (PFGE), has improved understanding of *M. agalactiae* CA at the national level. In particular, it was established that the various waves of ovine CA in the milk-producing region of the Pyrenees, including the most recent wave, were all the resurgence of a single clone located in this high livestock density region for at least the past 30 years (Nouvel *et al.,* 2012). In contrast, the strains of *M. agalactiae* isolated from sporadic caprine outbreaks are highly diversified, indicating a diffuse long-term enzootic in the country. More recently, strains of *M. agalactiae* isolated from Alpine ibex were found to be: i) very similar to one another, but ii) different from strains historically responsible for domestic CA in goats in the same valleys in Savoie, and iii) atypical compared to all currently known domestic strains, indicating an enzootic that is probably long-standing and specific to wild ungulates (Tardy *et al.,* 2012). In *Mmc* caprine CA, asymptomatic carriage and shedding appear to be frequent, with the outer ear in goats forming a favoured site where several strains or even species of mycoplasmas coexist (Mercier *et al.,* 2007). A series of surveys performed in partnership with ANSES Niort Laboratory provided an estimate of the prevalence of carrier level and allowed to collect strains that were not accessible *via* Vigimyc. In herds with no known history of mycoplasmosis, on average 8% of animals were *Mmc* carriers in the outer ear, and 5% of bulk tank milk was positive for *Mmc* (Tardy *et al.,* 2007). Subtyping of the various strains of *Mmc* by PFGE and micro-restriction followed by Southern Blot analysis of the insertion sequence profile showed, i) very high polymorphism in ear strains, ii) coexistence of several clones in healthy animals or in herds with no associated clinical signs, and, in contrast, iii) circulation of a single clone during a disease episode. Nonetheless, no difference was found between the carrier strains and clinical outbreak strains from Vigimyc, whether genetically or in terms of experimental virulence potential (Tardy *et al.,* 2010). As such, *Mmc* mycoplasmoses in goats appear to be latent enzootic infections with sporadic emergence of pathogenic strains. In these circumstances, applying a purely health-based prophylactic programme would seem bound to fail (Tardy *et al.,* 2007).

Given its current operational framework, Vigimyc is very effective at monitoring strain evolution from an antigenic and/or genetic point of view. Detailed characterisation of atypical strains is essential in order to maintain health watch that takes into account the genomic diversity of the various strains and their evolution, as well as emergence of new species or variants. Recently, a strain isolated from a clinical sample of caprine arthritis reacted with the specific monoclonal antibody targeting *Mmm*, the causative agent of CBPP. After molecular assessment, it was found to belong to the *Mmc* species. This cross-reaction was alarming bearing in mind that goats could be an occasional reservoir for CBPP, and could have cast doubt on the reliability of the serological screening test recommended for CBPP which is based on a competitive ELISA assay using the target epitope of this monoclonal antibody. A study of the variability of the epitope coding region in all the strains of *Mmc* enabled us to demonstrate that the probability of false positive clones is very rare and random (Tardy *et al.,* 2011) and therefore does not cast doubt on the reliability of screening, nor on the Vigimyc surveillance strategy in France.

Antimicrobial susceptibility

Unlike many other pathogenic bacteria in ruminants, mycoplasmas are not included in various surveillance networks for antimicrobial resistance since evaluating their antibacterial susceptibility requires specific techniques that are not available routinely in partner laboratories. However, reports from the field regularly indicate treatment failures with progression to chronic disease. In view of this, we used the strains available via Vigimyc to evaluate the current level of susceptibility of mycoplasma strains.

The first species tested was *M. bovis* which is often involved in infectious enzootic bronchopneumonia of calves, a multifactorial disease requiring large quantities of antibiotics and for which no thorough evaluation had been performed in France for 20 years. Using our strain collection, partly collected through Vigimyc, we were able to compare minimum inhibitory concentrations of various antibiotics used in veterinary medicine and likely to be active against mycoplasmas in 27 older isolates (1978-1979) and 46 recent isolates (2010-2012) of *M. bovis* from 73 separate outbreaks of infectious enzootic bronchopneumonia across France (Gautier-Bouchardon *et al.,*

Research

2014). A statistically representative loss of sensitivity was found for 8 antibiotics among 100% of contemporary strains. As a result, if we consider critical values accepted for pathogenic bacteria in the respiratory area in cattle, all contemporary mycoplasma strains would be classified as “resistant” to macrolides, tetracyclines, spectinomycin and florfenicol, and “intermediate” for fluoroquinolones. Bearing in mind that mycoplasmas are naturally resistant to all antibiotics that act on the cell wall (beta-lactams and glycopeptides), the therapeutic armamentarium against *M. bovis* mycoplasmoses would be extremely limited. Evaluation of the baseline level of resistance of other mycoplasma species to the various antibiotics used today is underway.

Use of strain diversity to develop knowledge on the Mycoplasma genus

The collection of strains obtained through Vigimyc provides biological resource for a number of research projects on the evolution of mycoplasmas, borders between species, virulence of strains, etc. In return, the information obtained is used to adjust our microbiological surveillance of mycoplasmoses. An example is the EVOLMYCO project (ANR-07-GMGE-001). It has provided the scientific community with 20 additional ruminant mycoplasma genome sequences, among which 8 correspond to strains from the Vigimyc collection (Dordet-Frisoni *et al.,* 2013; Dupuy *et al.,* 2013; Manso-Silvan *et al.,* 2013; Tardy *et al.,* 2012). Initial results from comparative genomics are forcing us to rethink current knowledge in mycoplasmology, and show that very few families of genes clearly distinguish strains based on their pathogenicity or their host. In addition, significant levels of horizontal gene transfer (HGT) between species that are not closely related but share the same ecological niche have been suggested *in silico* (Sirand-Pugnet *et al.,* 2007), calling into question the idea that mycoplasmas evolved primarily through downsizing, with massive gene losses. Our collection of strains has enabled us to look for potential vectors of HGT. Plasmids are minor contributors to HGT (Breton *et al.,* 2012), but integrative and conjugative elements (ICEs) appear far more promising and their inter-strain transfers have recently been reproduced *in vitro* (Dordet Frisoni *et al.,* 2013). Today, mycoplasmas appear to be genetic mosaics (Marenda, 2014). We have shown that the *M. leachii* species is in fact a genomic chimera between the *capricolum* and *mycoides* species and represents an excellent example of the genetic continuum between strains, beyond species borders (Tardy *et al.,* 2009). This new concept could call into question the very notion of species and thereby the taxonomy currently used to diagnose animal mycoplasmoses. In these conditions, diagnosis will probably move more toward an overall approach to mycoplasma diseases, with detection of trans-taxon virulence markers.

Table 1. Key figures for the Vigimyc Network for the 2009-2013 period: Number of treated outbreaks, studied animal species, types of animals and frequency of the various diseases of interest

	Hosts			
	Cattle	Goats	Sheep	Wildlifee
Volume of analyses and origin of strains				
Number of départements providing samples	62	55	34	1
Number of analysed isolates	856	725	237	120
Number of outbreaks	735	511	192	89
Distribution of samples based on host animal age (%)				
Adult animals	7	71	18	80
Young animals	76	18	57	5
Animal of unknown age	17	11	25	15
Distribution of samples based on type of disease (%) (Disease present alone or in combination with other clinical signs)				
Respiratory disease	89	24	68	32
Mastitis	2	36	2	0
Arthritis	2	15	2	0
Otitis	1	0	0	0
Septicaemia	0	2	0	0
Abortion	0	0	3	0
Eye disease	0	0	3	4
Unknown disease	5	19	20	21
Health follow-up	0	0	0	43
No disease	1	4	2	0

Research

Conclusion

The Vigimyc Network is unique in continental Europe, with only the UK having a comparable system. After a decade of operation, Vigimyc has largely fulfilled its initial objective of reviewing the epidemiological situation of regulated mycoplasmoses and mycoplasmoses with an economic impact in ruminants nationally. With its current organisation, characterised by an overall approach to all mycoplasmas and mycoplasmoses, it is perfectly suited to future changes in diagnosis and surveillance. Its strong point in the last few years has been above all the efforts to make the most of the strain collection generated by Vigimyc. This is probably the area in which Vigimyc will strengthen its activities by transferring its diagnostic role to partner laboratories, a change that has become possible thanks to recent technical innovations.

Acknowledgements

The authors would like to thank all Vigimyc partner laboratories, the Vigimyc technical team at ANSES and in particular Patrice Cuchet and Véronique Lefriand, as well as Jean-Luc Vinard for the design and development of the Vigimyc database.

Table 2: Distribution of the 2105 isolates (including mixes of species) identified between 2009 and 2013 based on animal species

	Host animal								
	Cattle (n=1029)		Goats (n=703)		Sheep (n=275)		Ibex (n=98)		Total
(sub)-species of mycoplasma	n	%	n	%	n	%	n	%	n
Pathogenic									
M. agalactiae	0		29		2		16	16.3	47
M. bovis	488	47.4	0		1		0		489
M. capricolum subsp. capricolum	0		192	27.3	0		1		193
M. capricolum subsp. capripneumoniae	0		0		0		0		0
M. leachii	0		0		0		0		0
M. mycoides subsp. capri	5		288	41	6		0		299
M. mycoides subsp. mycoides	0		0		0		0		0
M. putrefaciens	0		85		0		1		86
Unclear pathogenic potential									
M. alkalescens	34		0		0		0		34
M. canadense	13		0		0		0		13
M. canis	5		0		0		0		5
M. conjunctivae	0		0		6		0		6
M. feriruminatoris subsp. nov.	0		0		0		72	73.5	72
M. ovipneumoniae	0		24		78	28.4	0		102
Opportunistic									
Acholeplasma laidlawii	3		1		0		0		4
M. arginini	121		75		180	65.5	6		382
M. auris	0		3		0		2		5
M. bovigenitalium	10		1		2		0		13
M. bovirhinis	350	34	0		0		0		350
M. edwardii	0		1		0		0		1
M. yeatsii	0		4		0		0		4

n=number of isolates; %=proportion of isolates by host animal (the proportion is given only for the two most common (sub)-species).

Research

References

Arcangioli, M., Chazel, M., Sellal, E., Botrel, M., Bezille, P., Poumarat, F., Calavas, D., Le Grand, D., 2011. Prevalence of *Mycoplasma bovis* udder infection in dairy cattle: Preliminary field investigation in southeast France. New Zealand Veterinary Journal 59, 75-78.

Becker, C.A., Ramos, F., Sellal, E., Moine, S., Poumarat, F., Tardy, F., 2012. Development of a multiplex real-time PCR for contagious agalactia diagnosis in small ruminants. Journal of microbiological methods 90, 73-79.

Breton, M., Tardy, F., Dordet-Frisoni, E., Sagne, E., Mick, V., Renaudin, J., Sirand-Pugnet, P., Citti, C., Blanchard, A., 2012. Distribution and diversity of mycoplasma plasmids: lessons from cryptic genetic elements. BMC microbiology 12, 257.

Chazel, M., Tardy, F., Le Grand, D., Calavas, D., Poumarat, F., 2010. Mycoplasmoses of ruminants in France: recent data from the national surveillance network. BMC Veterinary Research 6, 32.

Dordet-Frisoni, E., Baranowski, E., Barre, A., Blanchard, A., Breton, M., Couture, C., Dupuy, V., Gaurivaud, P., Jacob, D., Lemaitre, C., Manso-Silvan, L., Nikolski, M., Nouvel, L.X., Poumarat, F., Sirand-Pugnet, P., Thebault, P., Theil, S., Thiaucourt, F., Citti, C., Tardy, F., 2013. Draft genome sequences of *Mycoplasma auris* and *Mycoplasma yeatsii*, two Species of the ear canal of caprinae. Genome announcements 1.

Dordet Frisoni, E., Marenda, M.S., Sagne, E., Nouvel, L.X., Guerillot, R., Glaser, P., Blanchard, A., Tardy, F., Sirand-Pugnet, P., Baranowski, E., Citti, C., 2013. ICEA of *Mycoplasma agalactiae*: a new family of self-transmissible integrative elements that confers conjugative properties to the recipient strain. Molecular microbiology 89, 1226-1239.

Dupuy, V., Sirand-Pugnet, P., Baranowski, E., Barre, A., Breton, M., Couture, C., Dordet-Frisoni, E., Gaurivaud, P., Jacob, D., Lemaitre, C., Manso-Silvan, L., Nikolski, M., Nouvel, L.X., Poumarat, F., Tardy, F., Thebault, P., Theil, S., Citti, C., Blanchard, A., Thiaucourt, F., 2013. Complete genome sequence of *Mycoplasma putrefaciens* strain 9231, one of the agents of contagious agalactia in goats. Genome announcements 1.

Gautier-Bouchardon, A.V., Ferre, S., Le Grand, D., Paoli, A., Gay, E., Poumarat, F., 2014. Overall decrease in the susceptibility of *Mycoplasma bovis* to antimicrobials over the past 30 years in France. PloS one 9, 1-9, e87672.

Le Grand, D., Saras, E., Blond, D., Solsona, M., Poumarat, F., 2004. Assessment of PCR for routine identification of species of the *Mycoplasma mycoides* cluster in ruminants. Veterinary Research 35, 635-649.

Maigre, L., Citti, C., Marenda, M., Poumarat, F., Tardy, F., 2008. Suppression-subtractive hybridization as a strategy to identify taxon-specific sequences within the *Mycoplasma mycoides* cluster: design and validation of an *M. capricolum* subsp. *capricolum*-specific PCR assay. Journal of Clinical Microbiology 46, 1307-1316.

Manso-Silvan, L., Tardy, F., Baranowski, E., Barre, A., Blanchard, A., Breton, M., Couture, C., Citti, C., Dordet-Frisoni, E., Dupuy, V., Gaurivaud, P., Jacob, D., Lemaitre, C., Nikolski, M., Nouvel, L.X., Poumarat, F., Thebault, P., Theil, S., Thiaucourt, F., Sirand-Pugnet, P., 2013. Draft genome sequences of *Mycoplasma alkalescens*, *Mycoplasma arginini*, and *Mycoplasma bovigenitalium*, three species with equivocal pathogenic status for cattle. Genome announcements 1.

Marenda, M. 2014. Genomic Mosaics, In: Browning, G.F., Citti, C. (Eds.) Mollicutes: molecular biology and pathogenesis. Caister Academic Press, Norfolk, UK, 2-15.

Marenda, M.S., Sagne, E., Poumarat, F., Citti, C., 2005. Suppression subtractive hybridization as a basis to assess *Mycoplasma agalactiae* and *Mycoplasma bovis* genomic diversity and species-specific sequences. Microbiology 151, 475-489.

Mercier, P., Pellet, M.P., Morignat, E., Calavas, D., Poumarat, F., 2007. Prevalence of mycoplasmas in external ear canal of goats: influence of the sanitary status of the herd. Small Ruminant Research 73, 296-299.

Nouvel, L.X., Marenda, M.S., Glew, M.D., Sagne, E., Giammarinaro, P., Tardy, F., Poumarat, F., Rosengarten, R., Citti, C., 2012. Molecular typing of *Mycoplasma agalactiae*: tracing european-wide genetic diversity and an endemic clonal population. Comparative Immunology, Microbiology and Infectious Diseases.

Pereyre, S., Tardy, F., Renaudin, H., Cauvin, E., Del Pra Netto Machado, L., Tricot, A., Benoit, F., Treilles, M., Bebear, C., 2013. Identification and subtyping of clinically relevant human and ruminant mycoplasmas by use of matrix-assisted laser desorption ionization-time of flight mass spectrometry. Journal of Clinical Microbiology 51, 3314-3323.

Poumarat, F., Perrin, B., Longchambon, D., 1991. Identification of ruminant mycoplasma by dot-immunobinding on membrane filtration (MF dot). Veterinary Microbiology 29, 329-338.

Sirand-Pugnet, P., Lartigue, C., Marenda, M., Jacob, D., Barré, A., Barbe, V., Schenowitz, C., Mangenot, S., Couloux, A., Segurens, B., de Daruvar, A., Blanchard, A., Citti, C., 2007. Being pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. PLoS Genetics 3, e75.

Tardy, F., Baranowski, E., Nouvel, L.X., Mick, V., Manso-Silvan, L., Thiaucourt, F., Thebault, P., Breton, M., Sirand-Pugnet, P., Blanchard, A., Garnier, A., Gibert, P., Game, Y., Poumarat, F., Citti, C., 2012. Emergence of atypical *Mycoplasma agalactiae* strains harbouring a new prophage and associated with a mortality episode of Alpine wild-ungulates. Applied and Environmental Microbiology 78, 4659-4668.

Tardy, F., Gaurivaud, P., Manso-Silvan, L., Thiaucourt, F., Pellet, M.P., Mercier, P., Le Grand, D., Poumarat, F., 2011. Extended surveillance for CBPP in a free country: challenges and solutions regarding the potential caprine reservoir. Preventive Veterinary Medecine 101, 89-95.

Tardy, F., Gaurivaud, P., Tricot, A., Maigre, L., Poumarat, F., 2008. Epidemiological surveillance of mycoplasmas belonging to the '*Mycoplasma mycoides*' cluster: is DGGE fingerprinting of 16S rRNA genes suitable? Letters in Applied Microbiology 48, 210-217.

Tardy, F., Maigre, L., Poumarat, F., Citti, C., 2009. Identification and distribution of genetic markers in three closely related taxa of the *Mycoplasma mycoides* cluster: refining the relative position and boundaries of the *Mycoplasma sp.* bovine group 7 taxon (*Mycoplasma leachii*). Microbiology 155, 3775-3787.

Tardy, F., Maigre, L., Tricot, A., Poumarat, F., Nguyen, L., Le Grand, D., 2010. Comparison of isolates of *Mycoplasma mycoides* subspecies *capri* from asymptomatic and septicaemic goats. Journal of Comparative Pathology 144, 70-77.

Tardy, F., Mercier, P., Solsona, M., Saras, E., Poumarat, F., 2007. *Mycoplasma mycoides* subsp. *mycoides* biotype large colony isolates from healthy and diseased goats: prevalence and typing. Veterinary Microbiology 121, 268-277.

Euro Reference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Research

Integrated veterinary networks for the surveillance of zoonotic agents in Italy

Antonia Ricci (aricci@izsvenezie.it) - OIE/National Reference Laboratory for Salmonella
Istituto Zooprofilattico Sperimentale delle Venezie, Viale dell'Università 10, I 35020 Legnaro (PD), Italie

Keywords: network, Italy, zoonotic pathogens.

Foodborne zoonoses are infections and diseases naturally transmissible between animals and humans via contaminated foodstuffs. The severity of these diseases in humans varies from subclinical infection or mild symptoms to life-threatening conditions. In order to prevent such zoonoses from occurring, it is important to identify which animals and foodstuffs are the main sources of infection. For this purpose, information aimed at protecting human health is collected and analysed from all European Union Member States. The European Union (EU) system for the monitoring and collection of information on zoonoses is based on the Zoonoses Directive 2003/99/EC, which obliges EU Member States to collect relevant and, where applicable, comparable data on zoonoses, zoonotic agents, antimicrobial resistance and foodborne outbreaks. According to the data reported in the latest European Union Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents and Food-borne Outbreaks, published by EFSA in 2014, *Salmonella* remained the most frequently reported cause of foodborne outbreaks in the EU in 2012 (1,533 of 5,363 foodborne outbreaks, 28.6%).

In 2012, a total of 91,034 confirmed cases of human salmonellosis were reported in the EU. This represents a decrease of 4.7% compared with 2011 and a decrease of 43,546 cases (32%) compared with the case numbers reported in 2008. The EU notification rate for confirmed cases was 22.2 cases per 100,000 population. The EU case-fatality rate was 0.14%, as 61 deaths due to non-typhoidal salmonellosis were reported in the EU in 2012. As in previous years, S. Enteritidis and S. Typhimurium were the most frequently reported serovars (41.3% and 22.1%, respectively, of all known reported serovars in human cases); monophasic S. Typhimurium 1,4,[5],12:i:- was the third most commonly reported serovar in the EU (7.2%). The fourth most common serovar in humans was *Salmonella* Infantis, with increasing numbers of reported isolates over the last five years (EFSA, 2014).

Knowledge about strain characteristics, through the typing of isolates, is of paramount importance in order to investigate sources of disease by comparing strains isolated from different reservoirs. The harmonised serotyping system is crucial in establishing the epidemiology of *Salmonella* at all levels. Ideally, a large proportion of, if not all, *Salmonella* isolates from humans, animals and food should be subtyped, to support epidemiological insight and to target interventions to prevent infections from identified sources (Wagenaar *et al.,* 2013). Information is needed about the presence of pathogens in the full array of exposure sources (e.g., foods, animals, drinking water, and recreational water) (Parmley *et al.,* 2013). Integrated surveillance programmes collect samples and generate information from multiple components within a system (Galanis *et al.,* 2012).

Microbial subtyping is one of the major methodologies to attribute food-borne infectious diseases to their sources. The principle behind source attribution is the comparison of subtypes of isolates causing human disease with the distribution of these subtypes in their putative sources (e.g. animals, food, or the environment) (Pires *et al.,* 2009). Microbial subtyping source attribution relies on laboratory subtyping methods to identify overlaps between subtypes identified from cases of human disease and those from their potential sources (Barco *et al.,* 2013).

Human infections caused by subtypes that have been exclusively or almost exclusively isolated from a single source can then be attributed to that specific source. On the other hand, when human infections are caused by subtypes isolated in several sources, they can be attributed to those sources proportionally to the reported occurrence of subtypes in the sources (Hald *et al.,* 2004).

In Italy, in order to collect data about *Salmonella* strains isolated in the veterinary sector, the Enter-Vet network was established in 2002, with the aim of collecting data at the national level on *Salmonella* spp. detection from samples of animal origin. This network operates in close collaboration with the Enter-Net system, which manages data on strains of human origin.

The Enter-Vet network consists of the laboratories of Istituti Zooprofilattici Sperimentali (IZS) under the supervision of the National Reference Laboratory for *Salmonella* (NRL). The Istituto Zooprofilattico Sperimentale delle Venezie (IZSVe) was designated as the National Reference Laboratory for *Salmonella* in 1999 by the Ministry of Health. In 2007, the laboratory was recognised by the World Organisation for Animal Health (OIE) as a Reference Laboratory for salmonellosis. The laboratories participating in the Enter-Vet network send to the NRL data on *Salmonella* spp. strains serotyped, together with *Salmonella* Enteritidis, *Salmonella* Typhimurium and the monophasic variant of *Salmonella* Typhimurium isolates to be phage typed. Serotyping is carried out by all laboratories of the network according to the Kauffman-White scheme, while phage typing is performed by the National Reference Laboratory according to the schemes provided by the Health Protection Agency (Colindale, London, UK).

In order to guarantee the quality of the data produced by the network, the NRL organises annual ring trials for the laboratories of the network, as external quality control for *Salmonella* serotyping and detection. The NRL publishes an annual report of Enter-Vet activity, which can be downloaded from the IZSVe website (www.izsvenezie.it). The reports give an overview on

Research

the distribution of *Salmonella* serotypes and phage-types in different animal species and food sources at the national level, providing information useful for targeting epidemiological investigations in cases of human outbreaks. Moreover, all the strains are properly stored as well as the related epidemiological information, making it possible to perform more accurate typing, also through molecular methods, on subsets of strains that appear to be relevant at a certain time.

In 2012, the network collected information on 3,567 *Salmonella* strains, 58% of them isolated from animals at primary production, 30% from food matrices, and 4% from feeding stuffs and environmental samples (mostly at farm level). Poultry and swine are the most represented animal species. **Table 1** reports the distribution of serotypes in the different sources.

Table 1: serovars isolated from different sources, Enter-Vet, 2012

Serovar	Animal	Food	Animal feed	Environment	Unknown	Water	Total	%
Monophasic variant S. Typhimurium	237	266	2	27	3	2	537	15.05
Typhimurium	199	92	3	6	10	5	315	8.83
Derby	95	166	2	5	0	2	270	7.57
Livingstone	144	19	15	8	2	1	189	5.30
Infantis	59	53	1	7	0	0	120	3.36
Rissen	37	71	3	1	0	1	113	3.17
Thompson	90	13	2	5	1	0	111	3.11
Enteritidis	59	11	0	19	2	0	91	2.55
Agona	43	6	32	6	2	1	90	2.52
Mbandaka	57	4	16	3	5	0	85	2.38
Bredeney	51	27	2	2	1	0	83	2.33
Kentucky	71	7	0	5	0	0	83	2.33
Hadar	35	23	0	1	0	4	63	1.77
Veneziana	52	5	0	1	0	5	63	1.77
Muenchen	27	23	0	9	0	1	60	1.68
London	31	23	1	1	2	0	58	1.63
Braenderup	48	3	1	2	2	1	57	1.60
Coeln	52	3	0	2	0	0	57	1.60
Newport	32	19	2	1	0	2	56	1.57
Choleraesuis	45	3	0	0	0	0	48	1.35
Blockley	30	11	1	0	5	0	47	1.32
Saintpaul	16	21	0	1	0	0	38	1.07
Give	10	16	0	3	0	1	30	0.84
Anatum	14	11	3	1	0	0	29	0.81
Other serovar	349	143	60	27	14	18	611	17.13
Other	200	43	9	6	1	4	263	7.37
Total	2083	1082	155	149	50	48	3567	100.00

Surveillance is critical to the policy-making process, providing evidence needed to target interventions to improve food and water safety, and ultimately reduce the burden of disease. Laboratory surveillance, of which the Enter-Vet network is an example, is one of the pieces of this puzzle, whose probability of success in identifying the sources of disease, and therefore applying effective control measures, is a function of the level of integration within the system.

References

Barco, L., Barrucci, F., Olsen, J.E., Ricci, A. 2013. *Salmonella* source attribution based on microbial subtyping. *Int J Food Microbiol*, 163: 193-203.

EFSA (European Food Safety Authority) and ECDC (European Centre for Disease Prevention and Control), 2014. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2012. *EFSA Journal*, 12: 3547, 312 pp.

Galanis, E., Parmley, J., DeWith, N., British Columbia integrated surveillance of foodborne pathogens working group. 2012. Integrated surveillance of *Salmonella* along the food chain using existing data and resources in British Columbia, Canada. *Food Res Int*, 45:795-801.

Hald, T., Vose, D., Wegener, H.C., Koupeev, T. 2004. A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. *Risk Anal* 24: 255-269.

Wagenaar, J.A., Hendriksen, R.S., Carrique-Mas, J. 2013. Practical considerations of surveillance of *Salmonella* serovars other than Enteritidis and Typhimurium. *Rev Sci Tech* OIE,32: 509-519.

Parmley, E. J., Pintar, K., Majowicz, S., Avery, B., Cook, A., Jokinen, C., Gannon, V., Lapen, D. R., Topp, E., Edge T.A., Gilmour, M., Pollari, F., Reid-Smith, R., and Irwin, R. 2013. A Canadian application of one health: integration of *Salmonella* data from various Canadian surveillance programs (2005–2010). *Foodborne Pathog Dis*, 10: 747-756.

Pires, S.M., Evers, E.G., van Pelt, W., Ayers, T., Scallan, E., Angulo, F.J., Havelaar, A., Hald, T. Med-Vet-Net Workpackage 28 Working Group. 2009. Attributing the human disease burden of foodborne infections to specific sources. *Foodborne Pathog Dis*, 6: 417-424.

Euro Reference
journal of Reference

anses
French agency for food, environmental and occupational health & safety

Summary | Point of view | Focus | Methods | Research | Agenda

Journal No. 12
Summer 2014

Research

Development of a European molecular typing database for food, environmental and veterinary *Listeria monocytogenes* strains

Benjamin Félix[1*] (benjamin.felix@ANSES.fr), Corinne Danan[1], Pia Makela[2], Ivo Van Walle[3], Renaud Lailler[1], Thomas Texier[4], Bertrand Lombard[1],
Anne Brisabois[1], Sophie Roussel[1]

1. University Paris-Est, ANSES, Laboratory for Food Safety, European Union Reference Laboratory for *Listeria monocytogenes*, Maisons-Alfort, France.
2. EFSA, Unit of the Biological Monitoring Unit, EFSA, European Food Safety Authority, Via Carlo Magno 1A43126 Parma, Italy.
3. ECDC, European Centre for Disease Prevention and Control, Food- and Waterborne Diseases and Zoonoses Programme, Tomtebodavägen 11 A, SE-17183 Stockholm, Sweden
4. University Paris-Est, ANSES, Information Technology Services, Maisons-Alfort, France

Keywords: PFGE, harmonization, typing database, surveillance, food chain.

Abstract

The European Union Reference Laboratory (EURL) for *Listeria monocytogenes* (*Lm*) is collaborating with a network of 35 European National Reference Laboratories (NRLs). These NRLs are responsible for typing *Lm* strains of food, feed and animals. The consolidation of NRL capacity for standardised typing resulted in the recent creation of a centralized molecular database. Data (typing results and epidemiologic information) are provided by each NRL and shared within the network. This database, together with databases on human strains, is to become part of the European surveillance system in order to improve the traceability of *Lm* strains circulating throughout Europe.

Introduction

In several European countries, after a long decline of human cases, the incidence of listeriosis has increased over the last decade (EFSA-ECDC, 2010; EFSA-ECDC, 2012; Goulet *et al,* 2008). Listeriosis, a consequence of the ingestion of *Listeria monocytogenes* (*Lm*), was reported in 26 European Union (EU) Member States (MSs) in 2012 (EFSA-ECDC, 2012).

Molecular typing of food-borne bacteria is an essential tool for various surveillance purposes, such as (1) monitoring the spread of clones and strains, (2) providing an essential tool for epidemiological investigations and early detection of scattered national or international outbreaks, and (3) predicting epidemic potential. For *Lm* typing, pulsed-field gel electrophoresis (PFGE) remains currently the "Gold Standard" method. Finding a PFGE profile of a strain isolated in food that matches a human strain profile does not necessary imply that this food is the source of the contamination. It could simply reflect circulation of this particular strain. Nevertheless detecting a human strain profile in food should improve the rapidity and precision of outbreak detection. PFGE profiles combined with epidemiological data must be considered together to conclude on case attribution. The goal is to collect as much scientific evidence as possible, to help decision makers to withdraw or to recall a product from the market.

In recent years, interest in developing a European surveillance network for listeriosis has led to enhanced surveillance activities in several countries and has generally heightened awareness of the public health importance of *Lm*. The surveillance network PulseNet Europe ceased to be active in 2006 by lack of funding (Swaminathan *et al,* 2006). However, efficient networks have been set up since and they cooperate closely to improve information exchange and molecular testing.

The ANSES Maisons-Alfort Laboratory for Food Safety has been designated European Union Reference Laboratory (EURL) for *Lm* (**http://www.ANSESpro.fr/eurl-listeria/**) by the Directorate-General for Heath and Consumers (DG Sanco) of the European Commission. It coordinates a network of 35 National Reference Laboratories (NRLs) in 29 Member States (MSs) and Norway. The majority of these NRLs are responsible for typing *Lm* strains of food, environmental and veterinary origin isolated nationally. Out of 35 NRLs, 6 NRLs are in charge of typing of clinical strains.

The European Food Safety Authority (EFSA) collects information from the Member States on food and animals as well as food-borne outbreaks on an annual basis. For this purpose, EFSA has created a Task Force on Zoonoses Data Collection with participants from all Member States, EEA countries, Switzerland, DG SANCO, and ECDC. For this data collection, EFSA runs a web-based reporting application as well as a more automated Data Collection Framework. The European Center for Disease Prevention and Control (ECDC) coordinates a network of national public health surveillance institutes and national public health laboratory (NPHLs), in particular responsible for typing *Lm* strains isolated from national clinical cases (ECDC). ECDC has also developed The European Surveillance System (TESSy) (van Walle, 2013) molecular surveillance database in the objective to be used to timely share molecular epidemiologic information and PFGE data from strains isolated from human cases and quickly recognize outbreaks at the European scale. At the European level, other than TESSy, there was no molecular database for centralizing and sharing molecular data obtained from food strains. For this reason, the EURL has recently set up a database for *Lm* that includes typing results as well as epidemiological information related to strains isolated from food, environmental or animal samples. This database, known as the "EURL *Lm* Database" (EURL *Lm* DB), is shared within the NRL network. Its objective is to compile data sets as complete as possible on *Lm* typing and epidemiological data to document the European food chain. This article describes in a first part the

Research

EURL typing activities which have led to the setting up of the EURL *Lm* DB, in a second part the different steps involved in developing the EURL *Lm* DB and lastly addresses an example of the use of this database.

Consolidation of NRL capabilities for standardized *Lm* typing

The EURL has developed NRLs standardized serotyping and PFGE methods for *Lm* that have been shared with all collaborating NRLs (Félix *et al,* 2012b). The EURL PFGE method was recently compared with the PulseNet USA method. The results were basically indistinguishable for both methods (Félix, personal communication). The EURL has developed a PFGE profile interpretation standard operating procedure (SOP) (Félix *et al,* 2012a) which reduces operator-dependent subjectivity. The EURL has also organized regular typing training sessions, annual workshops, and three typing proficiency testing (PT) trials (Félix *et al,* 2012b; Félix *et al,* 2013) for the NRLs over the past six years. The recent PT trial organized by the EURL and the ECDC focused on PFGE analysis and profiles interpretation according to the PFGE profile interpretation SOP (Félix *et al,* 2012a). Out of 28 participating laboratories, 16 including 10 NRLs and 6 NPHLs were considered as competent for PFGE typing of *Lm*. This PT trial provided a valuable opportunity to facilitate and to stimulate exchanges of reproducible PFGE profiles between human and food reference laboratories (Félix *et al,* 2013). All these activities (1) stimulate NRLs for typing (2) reinforce and consolidate NRL typing capabilities (3) harmonize *Lm* typing methods throughout Europe.

Development of the EURL *Lm* DB

Technical support

The EURL *Lm* DB is hosted by the EURL. A network management platform dedicated to molecular *Lm* typing data exchange was implemented (**https://moleculartyping-db.ANSES.fr/EUListNet**). Web networking is managed by a web server (BioNumerics (BN) Server Web Edition, version 6.1, Applied Maths, Sint-Martens-Latem, Belgium). The network is based on machine-to-machine communication over the Internet, allowing NRL databases to be shared with the EURL *Lm* DB. The EURL *Lm* DB was developed using modified PulseNet USA communication scripts and protocols with the agreement of the U.S. Center for Disease Control and Prevention (CDC). Modifications include use of the latest Web Edition of the BN Server, facilitating data exchange through Internet-based protocols. The various functionalities included in the database are described in **Figure 1**.

Organization

The EURL *Lm* DB steering committee (SCOM) comprise an equal number of representatives from eight participating NRLs and from the ECDC, EFSA and the EURL (administrator and curator of the EURL *Lm* DB).

The registration number of each strain is randomly generated by the BN server at submission (unique identification (UID) code of 33 alphabetic characters) and used as the central database identifier. The strain identification is given by two other information fields. The first field gives the identity of the submitting laboratory as XXYY. The code is composed of two characters (XX) given the national identity of the submitting NRL (e.g. IT is the ISO 3166-1-alpha-2 code used for Italy), the next two characters (YY) represent the submitting laboratory's national number (e.g. 01 would be the code for the first Italian laboratory subscribing to the EURL *Lm* DB project). The second information field is the initial strain number given by the submitting laboratory. Only the random anonymous key would be available to users in order to keep data owner identity anonymous.

Pulsotype nomenclature is defined according to the PulseNet USA pulsotype format (Gerner-Smidt *et al,* 2006) : pulsotypes are labeled with the "EU" tag, e.g. for a AscI profile: GX6A16.0001. EU, GX6 meaning *Lm*, A16 meaning the restriction enzyme AscI, 0001 the pulsotype number and EU the European tag. Each pulsotype is associated to information on its frequency within the whole database.

Epidemiological classification

The epidemiological data are recorded in accordance with a detailed epidemiological classification (**Figure 2**) which consists of several consecutive fields associated with predefined pick lists in the software. The epidemiologic classification structure was based on the dataset required in the EFSA epidemiological reporting system (EFSA, 2012). However, to simplify reporting, epidemiological information in the EURL *Lm* DB is restricted to food classification typically used in *Lm* risk assessments.

Data management

The EURL is the curator and administrator of the EURL *Lm* DB. The administrator manages the participant connections (logins and passwords), communication scripts and is responsible for EURL *Lm* DB maintenance (**Figure 1**).

The curator is in charge of validating each new profile submitted. The curator can directly modify the gel image processing parameters, and the profile. Each profile is analysed and identified according to the EURL PFGE profile interpretation method using an identification group-based SOP. The technical skills of the curator for PFGE gel interpretation are regularly updated in an internal assessment process that establishes the test for curator qualification. After curation, the curator designates profiles as either "confirmed" or "unsatisfactory".

Any change made by the curator is traced in the EURL *Lm* DB and visible to the user. These changes are automatically implemented in the local NRL database. The update process is automatic and can be activated by the NRL at any time. Script functionality allows the NRL to generate a report on its own database listing all the changes made by the curator on a given profile.

Participants

Criteria for joining the EURL *Lm* DB include (1) the successful participation in the most recent EURL PFGE and PFGE profile interpretation PT trials and (2) the availability of PFGE analysis software (BioNumerics version 6.6 or higher) equipped with a specific BN server communication script.

Participants must comply with the memorandum of understanding (MoU) that regulates EURL *Lm* DB use. The first item of the MoU stipulates that, by submitting data, all participants accept disclosure of the data to the other participating NRLs and the EURL. It also specifies that NRLs take part in the database on a voluntary basis and that they are responsible for the content of the data submitted. The NRLs retain ownership of their data and are free to publish their own PFGE profiles even if they have been submitted to the EURL *Lm* DB. The MoU also establishes that disclosure of EURL *Lm* DB data by the NRLs or the EURL requires a joint written

Research

agreement between partners. EURL *Lm* DB citations are only permitted if they do not disclose any specific data other than those owned by the disclosing party. The MoU also establishes the information routes in case of EURL *Lm* DB solicitation by DG SANCO, MSs or external bodies. Finally, the MoU specifies that NRLs complying with the MoU can consult the EURL *Lm* DB without restriction.

The NRLs are trained to the use of the EURL *Lm* DB through individual training sessions, scheduled at the EURL, on site, by phone, or by video. Moreover, individual technical follow-up and technical assistance by the EURL are available for each NRL.

Data consultation and submission (Figures 1 and 2)

Submission of data to the central database

These submitted data included PFGE profiles, serotype (molecular or conventional serotype) data and epidemiological information (see Figure 2). For PFGE data, participants have to submit *ApaI* and *AscI* profiles simultaneously.

Consultation of data of the central database

Users can consult the EURL *Lm* DB in two ways: (1) by matching their own profile against existing entries in the central database or (2) by consulting the database according to queries from epidemiological and typing data. In both cases, profiles obtained following database consultation may be downloaded from the EURL *Lm* DB to the user's local database. These downloaded profiles are available as profiles of the local database for statistical analysis and comparison until the active session of the consultation software is closed.

Benefits of the EURL *Lm* DB use by the NRLs

The NRLs can use the EURL *Lm* DB to interpret a PFGE profile in their own database. All PFGE profiles available in the EUR *Lm* DB can be compared with local NRL profiles. The EURL *Lm* DB can thus be considered as a routine assessment tool for NRLs. The use of EURL *Lm* DB facilitates the harmonisation of the PFGE profiles in the local NRL database.

For a given PFGE profile, the NRL can access following information available in the EURL *Lm* DB: (1) serotype, (2) food matrix, (3) sampling date and (4) profile frequency in the EURL *Lm* DB. The use of the EURL *Lm* DB thus enables NRLs to collect information useful for epidemiological investigation in case of an outbreak.

The database also could accept data resulting from alternative typing techniques such as multi-locus sequence typing (MLST), multi-locus variable number tandem repeats analysis (MLVA), and whole genome sequencing data. Data generated using these methods can be compared with PFGE and serotyping data to determine their congruity and the structure of profile groups.

Example of use of the EURL *Lm* DB

The EFSA set up a monitoring program (baseline survey) on the prevalence of *Lm* in selected categories of ready-to-eat (RTE) foods (Decision 2010/678/EU) in 2010-2011 in EU Member States (MSs). This survey should allow the comparison of *Lm* contamination in ready-to-eat food in the Community and Member States and the verification of the Community food safety criteria for *Lm*. The ECDC and the EURL *Lm*, in collaboration with EFSA and with approval of the European Commission and the EU MSs, have launched a joint project to compare the PFGE profiles of food isolates from the coordinated European monitoring programme with strains from human listeriosis cases isolated during the same period. At the ECDC, the European *Listeria* Typing Exercise (ELiTE) project focuses on PFGE typing of human *Listeria* strains that have been isolated and stored by NPHLs in 2010-2011. The EURL *Lm* coordinates PFGE typing of the strains collected by NRLs and isolated from certain RTE food categories sold at retail stage in EU MSs in 2010-201. NRLs, trained by the EURL and having demonstrated their competence through participation in inter-laboratory proficiency trials organised by the EURL, type these strains isolated on the national level using serotyping and PFGE methods. Likewise, the EURL types strains from other countries. Human and food typing data are compared to allow a better estimation of the importance of certain foods as sources of human listeriosis.

As part of the European baseline survey, nine NRLs currently use the EURL *Lm* DB to submit and share the PFGE profiles of strains isolated in certain RTE foods at the national level. The EURL also contributes the typing and epidemiological data on strains sent by NRLs of 16 other countries to the EURL *Lm* DB. The close collaboration between the ECDC and the EURL will foster the exchange and comparison of typing data, as part of the ECDC's ELiTE project. The SOP and PFGE profile nomenclature used for the ELiTE project and the TESSy pilot study are currently being discussed by the ECDC, the EURL for *Lm* and EFSA, taking into account those used for the EURL *Lm* DB.

Conclusion - Prospects

The use of the EURL *Lm* DB should encourage individual countries to strengthen national surveillance of *Lm* infections, by facilitating the implementation and wide use of the typing national databases in each country. The functionalities of EURL *Lm* DB associated with curation work make it easier to harmonise PFGE profiles and the epidemiological data circulating within the NRL/EURL database network.

It is expected that the EURL *Lm* DB will be used together with databases on human strains and by microbiologists and epidemiologists. Used in combination with collaborative epidemiological investigations, the EURL *Lm* DB should improve the surveillance of *Lm* in the food chain by (1) enhancing the detection of European, eventually transboundary, contamination clusters, (2) optimising the detection of emerging *Lm* strains potentially pathogenic for consumers, (3) facilitating communication between NRLs, the EURL, EFSA and the ECDC and (4) suggesting links with potential sources of contamination.

Acknowledgments

This work was conducted in the framework of the European Union Reference Laboratory for *Listeria monocytogenes* activities and was supported by a grant from the European Commission's Directorate General for Health and Consumer Protection (DG SANCO). We would like to thank the ECDC and EFSA for their participation in the project. We would also like to thank Peter Gerner-Smidt (CDC) and Eva Moller-Nielsen (Staten Serum Institute) for their willingness to share PN customized scripts and protocols. Finally, our thanks go to the study group composed of steering committee members and all the European NRLs involved through consultation of the network. We also thank the team of Peter Gerner-Smidt (CDC) and Brenda Brown for their helpful advises in the writing of this publication.

Research

References

ECDC http://ecdc.europa.eu/en/activities/surveillance/strategies_principles/Pages/case_definitions.aspx

EFSA-ECDC (2010) The Community Summary Report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in the European Union in 2008. *EFSA Journal* 2010 **8(1):1496**: 137-165

EFSA-ECDC (2012) The European Union Summary Report on Trends and Sources of Zoonoses, Zoonotic Agents and Food-borne Outbreaks in 2010. *EFSA Journal* 2012;**10(3):2597**: 133-160

EFSA. (2012) Manual for reporting of food-borne outbreaks in accordance with Directive 2003/99/EC from the year 2011.

Félix B, Brisabois A, Dao TT, Lombard B, Asséré A, Roussel S (2012a) The use of Pulsed Field Gel Electrophoresis in *Listeria monocytogenes* sub-typing : harmonization at the European Union level. In *Gel Electrophoresis : Principles and Basics*, INTECH (ed), Vol. 1, Dr. Sameh Magdeldin edn, 14, pp 241-254. Niigata University, Japan

Félix B, Dao TT, Grout J, Lombard B, Assere A, Brisabois A, Roussel S (2012b) Pulsed-Field Gel Electrophoresis, Conventional, and Molecular Serotyping of *Listeria monocytogenes* from Food Proficiency Testing Trials Toward an Harmonization of Subtyping at European Level. *Foodborne Pathog Dis* **9**: 719-726

Félix B, Vingadassalon N, Dao TT, Asséré A, Lombard B, Brisabois A, Roussel S (2013) PFGE Proficiency Testing Trials: toward an harmonization of typing of *Listeria monocytogenes* food and clinical strains at European Level. *Foodborne Pathog Dis* In press

Gerner-Smidt P, Hise K, Kincaid J, Hunter S, Rolando S, Hyytia-Trees E, Ribot EM, Swaminathan B (2006) PulseNet USA: a five-year update. *Foodborne Pathog Dis* **3**: 9-19

Goulet V, Hedberg C, Le Monnier A, de Valk H (2008) Increasing incidence of listeriosis in France and other European countries. *Emerg Infect Dis* **14**: 734-740

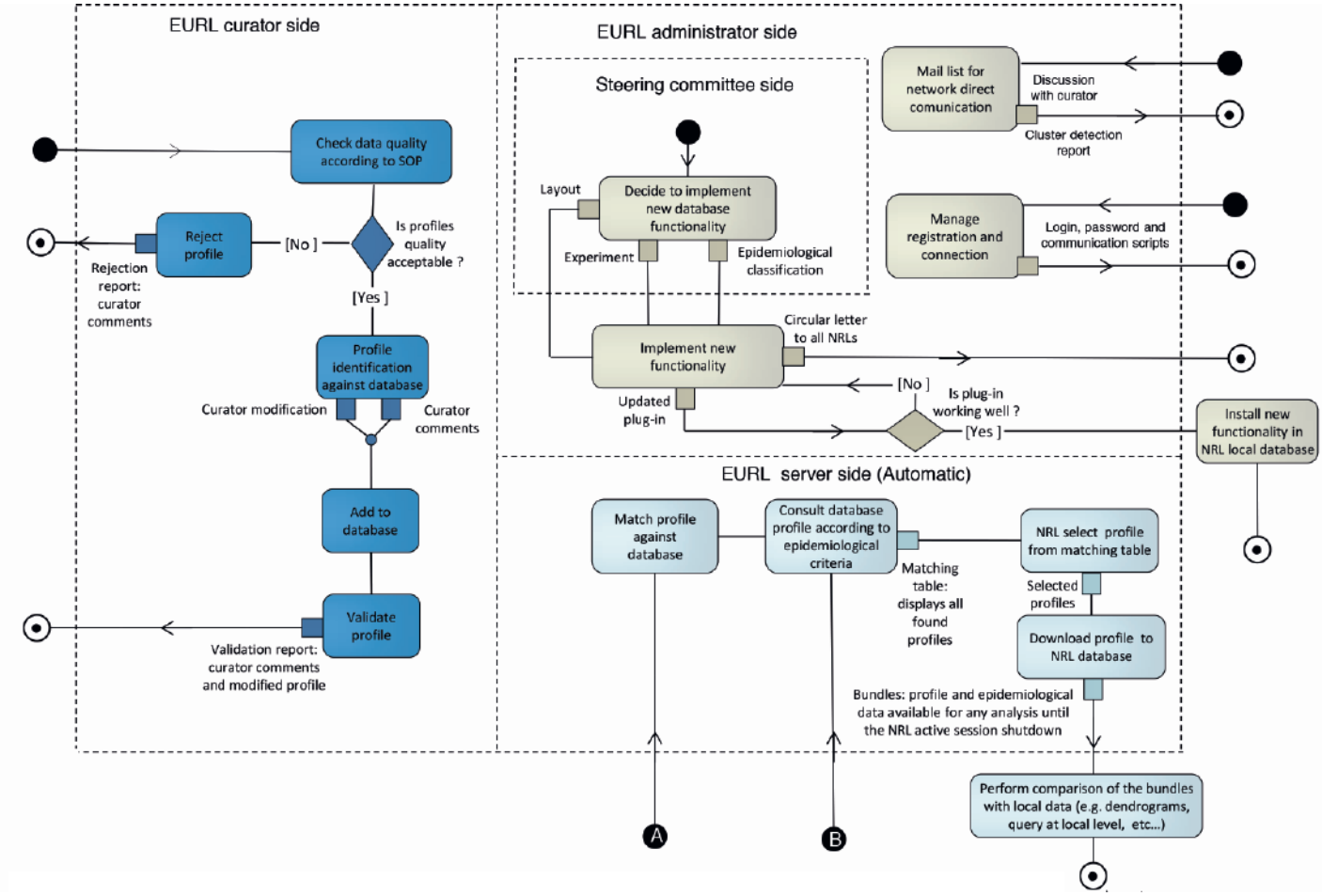
Graves LM, Swaminathan B (2001) PulseNet standardized protocol for subtyping *Listeria monocytogenes* by macrorestriction and pulsed-field gel electrophoresis. *Int J Food Microbiol* **65**: 55-62

Swaminathan B, Gerner-Smidt P, Ng LK, Lukinmaa S, Kam KM, Rolando S, Gutierrez EP, Binsztein N (2006) Building PulseNet International: an interconnected system of laboratory networks to facilitate timely public health recognition and response to foodborne disease outbreaks and emerging foodborne diseases. *Foodborne Pathog Dis* **3**: 36-50

van Walle I (2013) ECDC starts pilot phase for collection of molecular typing data. *Euro Surveill* **18 (3)**

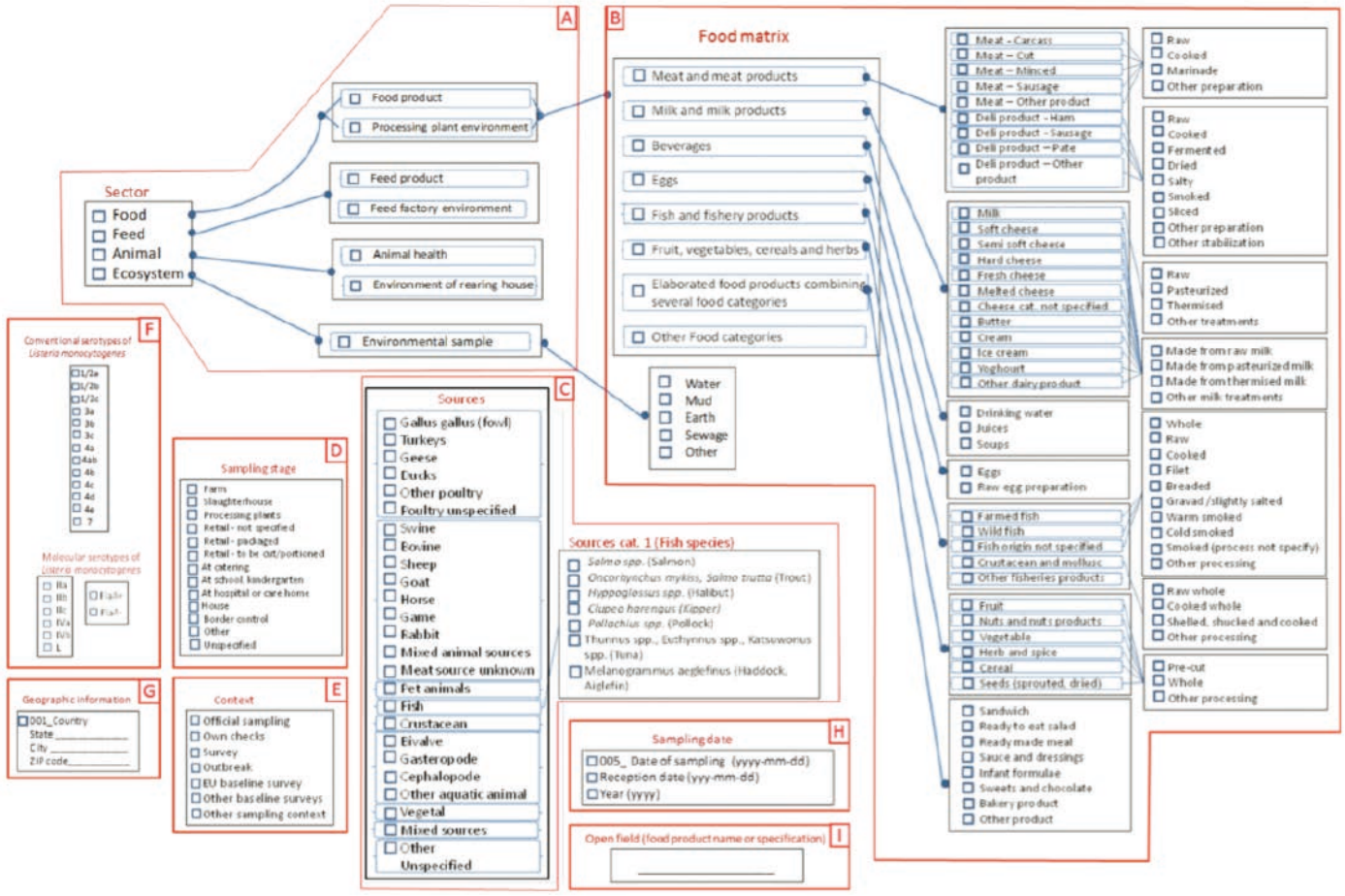
Research

Figure 1. Schematic representation of the EURL *Lm* DB data processing for administrator side (red), database automatic consultation (clear blue) and data submission (deep blue). Black circles symbolize when processes start and circles with dot when processes end up. The black circle marked with A shows the way for requesting the database based on molecular profiles, the black circle marked with B shows the way for requesting the database based on epidemiological information including serotyping data.



Research

Figure 2. Description of the samples submitted to the EURL *Lm* DB, according to seven main information blocks (A, B, C, D, E, F, G, H, I). Block A describes the sample (Food, Feed, Animal or Environment) and is further subdivided into subcategories describing where and how the sample was taken. Block B describes the food matrix, with first eight large categories of food products (Meat and meat products, Fish products, Elaborated food products combining several food categories or other food products, etc.) and further, specific information describing in detail the product and its processing. Block C describes the type of food sources which compose the product. Block D describes the level in the food chain where the sample was taken (from Farm to Retail or Borders). Block E describes the context of sampling, block H the date and block G the geographic information. Block F is composed of molecular or conventional serotyping data. Bloc I is an open field for the reporting of additional information.



EuroReference
journal of Reference
anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Research

The establishment of databases on circulating genotypes of *Mycobacterium tuberculosis* complex and web tools for an effective response to better monitor, understand and control the tuberculosis epidemic worldwide

David Couvin et Nalin Rastogi (nrastogi@pasteur-guadeloupe.fr)
WHO Supranational TB Reference Laboratory, Institut Pasteur de la Guadeloupe, 97183 Abymes, Guadeloupe, France.

Key words: *Mycobacterium tuberculosis*, tuberculosis, genotyping, databases, spoligotyping, MIRU-VNTR, epidemiology, phylogeny, drug-resistance.

Abstract

In this paper we will briefly review various *Mycobacterium tuberculosis* complex (MTBC) genotyping databases developed over last fifteen years at the Institut Pasteur de la Guadeloupe, which represent a great concerted initiative and effort to control the tuberculosis epidemic. Starting from the initial excel-based version in 1999 (SpolDB1; n=610 clinical isolates) to the fourth MySQL-based version in 2006 (SpolDB4; n=39,295 clinical isolates), these databases permitted to have a first phylogeographical snapshot of circulating MTBC genotypic lineages based on spacer oligonucleotide typing (spoligotyping) which allows to study the polymorphism of the Direct Repeat (DR) locus. The two most recent MySQL-based multimarker versions concern the 5th version SITVITWEB released in 2012 (n=62,582 clinical isolates) with both spoligotyping and 12-loci Mycobacterial Interspersed Repetitive Units - Variable-Number of Tandem Repeats (MIRU-VNTRs); and the 6th version named SITVIT2 that will be released in 2014 (n= 111,635 clinical isolates) with spoligotyping and 12-, 15- or 24-loci MIRU-VNTR data. In these recent versions, a web-based interface allows the user to search for strains through the database by criteria, such as the year, the isolation country, the country of origin, the investigator's name. It further facilitates to perform combined searches in SITVIT2, making it possible to get the genotyping data on selected strains in conjunction with their geographical distribution, as well as available data on drug-resistance, demographic and epidemiologic characteristics. Our research initiative is thus focused to further improve in depth phylogenetic characterization of MTBC lineages in conjunction with epidemiological analysis of circulating clones to generate evidence-based geographical mapping of predominant clinical isolates of tubercle bacilli causing the bulk of the disease both at country and regional level. Further superimposition of these maps with socio-political, economical, and demographical characteristics available through Geographic Information Systems (GIS) allows to have a precise view of prevailing disparities as seen at the level of United Nation's sub-regional stratification. An in-depth comprehension of these disparities and drawbacks is important to take appropriate actions by decision-makers and public health authorities alike, in order to better monitor, understand and control the tuberculosis epidemic worldwide.

Introduction

Almost twenty years after the World Health Organization (WHO) declaration of tuberculosis (TB) as a global public health emergency, and despite the major progress made towards 2015 global targets set within the context of the millennium development goals, TB is still the second deadliest disease caused by an infectious agent in the world after HIV/AIDS (WHO Report, 2013). In 2012, it led to an estimated 8.6 million new cases and 1.3 million deaths (including 320,000 deaths among TB/HIV co infected patients). A careful scrutiny of the WHO Report shows that TB remains an enormous health and economic problem not only in developing countries but also in developed nations due to the TB-HIV co-infection and emergence of multidrug-resistant (MDR), and more recently of extremely drug-resistant (XDR) isolates, which further complicates management of the disease and considerably increases the mortality due to TB among immune-compromised patients.

Furthermore, the increasing rate of travel/migration of population for leisure or work in the last decades has led to a new challenge in countries where TB was declining due to changes in the socio-epidemiological scenarios generated by massive immigration from countries where TB is highly endemic (García de Viedma *et al.,* 2011). Some key questions include comparison of the role of recent transmission with that of reactivation/importation in TB among foreign-born cases, the impact of potential importation of previously unidentified *M. tuberculosis* strains, and cross-transmission between cases from different nationalities. It is therefore important to understand how tubercle bacilli are transmitted, which clones are involved in drug-resistant cases and/or outbreaks, identify new clones that may be emerging in a setting vs. those that may be under extinction, identify subpopulations and risk factors with highest threat of catching the infection, and be able to interpret these results within the evolutionary framework of *M. tuberculosis* complex (MTBC). It is clear today that these questions would prove difficult to answer without the support of molecular epidemiology.

Until recently, all human-adapted strains of MTBC were traditionally considered to be essentially identical, hence the question of individual genetic variation within MTBC gained little attention which led to most previous research being focused on the individual organism. The advent of molecular methods and their widespread use in population-based studies introduced both new conceptual and technological developments. Although, MTBC constitutes a remarkably homogeneous group genetically with proven evidence of clonal evolution, recent studies have shown that the genetic diversity among individual clones is much higher than previously assumed, with potential impact on pathobiological properties. A leading study on a global collection of MTBC strains using seven megabase pairs of DNA sequence data showed significant genetic diversity

Research

(Hershberg *et al.,* 2008). The authors suggested that much of this genetic diversity was driven by genetic drift potentially linked to human demographic and migratory events with functional consequences such as emergence and spread of drug-resistant tuberculosis.
Although TB is present worldwide, some countries/regions contain higher incidences than others: e.g. Sub-Saharan Africa, Southern and Southeastern Asia, Latin America, but also Eastern Europe and Russia, as well as the Caribbean (particularly Haiti and the Dominican Republic). Indeed, TB still constitutes a major health problem in many of the Caribbean and Latin-American countries; e.g., the incidence of TB in Haiti which was 330/100,000 in the 90s is still as high as 213/100,000 in 2012 (WHO Report, 2013). It is obvious that early diagnosis and bacteriological confirmation of TB, and drug-resistance determination is a key process to control the TB epidemic, and reference laboratories represent essential structures in the global diagnosis and control of the disease worldwide. In such a context, Institut Pasteur took the initiative to start the first TB reference laboratory for the Caribbean region in 1993. Situated at Institut Pasteur de la Guadeloupe (IPG), it was initially targeted to work for the Caribbean; nonetheless being well aware of TB's global context, we started right from the beginning with an approach based on concomitant use of bacteriological and molecular methods. Over the last 2 decades, we developed a comprehensive approach which deals with routine and molecular diagnostics of MTBC and other mycobacterial species, drug-resistance surveillance, and development of rapid diagnostic methods and diverse molecular techniques that are useful for epidemiological and population-based studies at local, regional, and global level. For global TB surveillance, we developed a series of TB genotyping databases since 1999 which accumulated our own data as well as collected from various participating laboratories worldwide. This report will briefly summarize the steps undertaken to develop these databases and the web-based tools developed to offer the possibility of creating worldwide geographical distribution maps displaying the frequencies of TB genotypes worldwide at various geographical scales. We will also briefly refer to some of the published or ongoing studies making use of this information and future prospects.

TB: origin, spread and co-adaptation with its hosts

With evidence of the isolation and characterization of ancient *M. tuberculosis* DNA from an extinct bison dated 17,000 years B.C., suggesting the presence of TB in America in the late Pleistocene (Rothschild *et al.,* 2001), TB was already known as a very ancient disease. When looking at human remains, ancient DNA helped to trace the presence of TB in Egyptian mummies, with characterization of *M. tuberculosis* and *M. africanum* (Zink *et al.,* 2003). Subsequently, in analogy to other crowd diseases, the origin of human tuberculosis was thought to be associated with the Neolithic Demographic Transition (NDT) starting around 11,000 years ago, as the development of animal domestication increased the likelihood of zoonotic transfer of novel pathogens to humans while agricultural innovations supported increased population densities that helped sustain the infectious cycle (Wolfe *et al.,* 2007). However, it was not clear whether (a) human TB descended from a ruminant mycobacterium that recently infected humans from domestic animals, or from an ancient human mycobacterium that has come to infect domestic and wild ruminants; and (b) whether tuberculosis originated independently in both hemispheres or was brought to the Americas by Europeans. Nonetheless, TB also displayed a pattern of chronic progression, latency and reactivation which is characteristic of a pre-NDT disease (Barry *et al.,* 2009).
The answer to these questions came thanks to new research which showed that TB is probably as old as humanity itself (Comas *et al.,* 2013). By studying the diverse genetic variations in MTBC, the researchers were able to show that TB must have spread around the world with the first modern humans to emerge from Africa. This study analyzed the whole genomes of a collection of 259 contemporary strains of MTBC from around the world, and compared MTBC phylogenic diversity to human diversity inferred from mitochondrial genome data. The results indicated that MTBC emerged about 70,000 years ago, accompanied migrations of anatomically modern humans out of Africa, and expanded as a consequence of increases in human population density during the Neolithic period (Comas *et al.,* 2013). This early origin of TB and the fact that genome-based phylogeny of MTBC mirrored that of human mitochondrial genomes, further showed that TB lung infection did not spread to humans from their domesticated animals since farming came much later. This landmark study also showed striking similarities in the evolutionary path of humans and MTBC, and suggested that MTBC evolution not only paralleled that of humans but also that MTBC diversity directly benefited from human demographic explosions. The fact that *M. tuberculosis* "stricto-sensu" is an obligate human pathogen with no known animal or environmental reservoir, changes in human demography and population densities over time are most likely to affect its evolution like in case of a crowd disease model. At the same time, its latency and chronicity possibly allows it to adapt to lower host densities, survive, and strike back when favorable conditions allow massive host infections.

Typing methods for TB molecular epidemiology

MTBC is a very diverse group of organisms ecologically, and includes *M. tuberculosis, M. africanum*, and *M. canettii* (exclusively human pathogens), *M. microti* (a rodent pathogen), and *M. bovis* (bovine pathogen but with a wide host range), as well as *M. pinnipedii* (seals), *M. caprae* (goats), *M. mungi* (banded Mongooses), and the oryx (renamed *M. orygis*), dassie, and chimpanzee bacilli, causative agents of TB in the animal species after which they are named. However, MTBC constitutes a highly homogeneous group genetically and various MTBC members share on average more than 99.7 % of nucleotide identity (Kato-Maeda *et al.,* 2001). Despite this remarkable genetic homogeneity, the last 2 decades have witnessed the advent of new molecular methods widely used today in population-based genotyping studies, permitting to precisely characterize TB isolates, and infer different phylogenetic lineages associated. Detailed reviews are available on MTBC molecular evolution (Rastogi and Sola, 2007), current molecular typing methods (Jagielski *et al.,* 2014), and strategies and innovations in the broad field of TB molecular epidemiology (2) [García de Viedma *et al.,* 2011), and interested readers are referred to these for detailed information.
Although the IS6110-RFLP method was long considered the gold standard technique for *M. tuberculosis* typing due to its reproducibility and discriminatory power in the molecular epidemiological investigations of TB (van Embden *et al.,* 1993),

Research

this labor-intensive methodology required large quantities of DNA, and was characterized by a lack of discriminatory power for typing of the isolates with low copy numbers of IS6110, e.g. in South India (Radhakrishnan *et al.,* 2001). Furthermore, the fast molecular clock of this marker for evolutionary studies (de Boer *et al.,* 1999), the complexity of forces driving its transposition and risk of genetic convergence (Fang *et al.,* 2001), and difficulty to build large RFLP databases and need for sophisticated software for data analysis (Heersma *et al.,* 1998; Salamon *et al.,* 1998), rendered its use in evolutionary genetics of limited interest. For the reasons mentioned above, alternative PCR-based typing strategies such as Spoligotyping (Kamerbeek *et al.,* 1997) and mycobacterial interspersed repetitive units-variable number of tandem repeats or MIRU-VNTRs (Supply *et al.,* 2001; 2006) largely replaced IS6110-RFLP in the last decade; thus creating the basis for large-scale, high-throughput *M. tuberculosis* genotyping.

Based on the polymorphism of the direct repeat (DR) locus, Spoligotyping for "**SP**acer **OLIGO** nucleotide **TYPING**" is currently one of the most frequently used PCR-based approaches for studying the molecular epidemiology and phylogeography of MTBC. Initially identified in the vaccine strain M. bovis BCG (Hermans *et al.,* 1991), the DR locus belongs to the CRISPR ("clustered regularly interspaced short palindromic repeat") family of repetitive DNAs, and contains multiple identical 36-bp perfect repeats (3 helix turns) interspersed by unique 34-41 bp spacers. Together with their non-repetitive spacer sequences, the DR units constitute multiple direct variant repeats (DVRs) which show extensive polymorphism among *M. tuberculosis* clinical isolates. The 43 character binary pattern generated by this technique was previously shown to vehicle significant phylogenetic information (Sola *et al.,* 2001). Spoligotype patterns are now commonly designated with their octal description, an internationally-agreed reporting format (Dale *et al.,* 2001). Because of its simplicity, binary result format and high reproducibility, spoligotyping is widely used for investigations on MTBC molecular epidemiology as a macroarray-based method to study presence/absence of 43 selected spacers (out of 104 spacers present). Indeed, a PubMed search for "spoligotyping OR spoligotype" gave 924 published articles (interrogation made on 18 March 2014).

One of the limitations of Spoligotyping being its tendency to overestimate MTBC clustering, it was soon proposed to complement this method with a minisatellite (MIRU-VNTR) based typing in a «two PCR-based» strategy, in conjunction to conventional epidemiological investigations (Sola *et al.,* 2003). The MIRU-VNTR minisatellites constitute a multi-locus marker set since they represent independent markers of a same type, and have been used as classical 12-locus, discriminatory 15-locus, or full 24-locus formats (Supply *et al.,* 2001; 2006). However, even the 24-locus MIRUs lacked a satisfactory resolution power for accurately discriminating closely related Beijing genotype strains, a fact that led to a recent proposal to use an additional 4-locus set of consensus hypervariable MIRU-VNTRs for subtyping Beijing clonal complexes and clusters (Allix-Béguec *et al.,* 2014).

Molecular typing methods for TB phylogeny

Apart from their use in epidemiology (García de Viedma *et al.,* 2011), molecular typing methods are also useful for evolutionary studies (Rastogi and Sola, 2007). These essentially include 2 sets of markers: (a) those summarized above which are extensively used for epidemiological studies but also provide with concomitant phylogenetical information – IS6110-RFLP, spoligotyping, and MIRU-VNTRs (see below); and (b) a set of markers including Large Sequence Polymorphisms (LSP) / Regions of Difference (RD), and Single Nucleotide Polymorphisms (SNPs), that are specifically useful for phylogenetical and evolutionary studies.

One of the earliest studies used substractive genomic hybridization to identify three distinct genomic regions between virulent *M. bovis, M. tuberculosis*, and the avirulent *M. bovis* BCG strain, designated respectively as RD1, RD2, and RD3 (Mahairas *et al.,* 1996). In another study, a distinction between three genetic groups of *M. tuberculosis* was achieved based on two polymorphisms occurring at high frequency in the genes encoding catalase-peroxidase and the A subunit of gyrase, which led to a classification in three principal genetic groups (PGG); group 1 bacteria being ancestral to groups 2 and 3 (Sreevatsan *et al.,* 1997). Almost immediately thereafter, restriction-digested bacterial artificial chromosome (BAC) arrays of H37Rv strain were used to reveal the presence of 10 regions of difference between *M. tuberculosis* and *M. bovis* (RD1 to 10); 7 of which (RD4-RD10) were deleted in *M. bovis* (Gordon *et al.,* 1999). In a major contribution, Brosch *et al.* (2002) analyzed the distribution of 20 variable regions resulting from insertion-deletion events in the genome of the tubercle bacilli in a collection of strains belonging to all MTBC subspecies, and showed that based on the presence or absence of a M. tuberculosis specific deletion 1 (TbD1, a 2 kb sequence), M. tuberculosis could be divided into "ancient" TbD1 positive and "modern" TbD1 negative strains (Brosch *et al.,* 2002). In this new evolutionary scenario of the *M. tuberculosis* complex, the RD9 deletion identifies an evolutionary lineage represented by *M. africanum, M. microti* and *M. bovis* that diverged from the progenitor of the present *M. tuberculosis* strains before TbD1 occurred, a finding which contradicts previous assumptions that *M. tuberculosis* evolved from a precursor of *M. bovis* (Brosch *et al.,* 2002). Since *M. canettii* and other ancestral *M. tuberculosis* complex strains lacked none of these regions, they are supposed to be direct descendants of the tubercle bacilli that existed before the "*M. africanum - M. bovis*" lineage separated from the *M. tuberculosis* lineage.

Using a global MTBC collection and 212 SNPs, Filliol *et al.* (2006) identified six deeply branching, phylogenetically distinct SNP cluster groups (SCGs) and five subgroups. The SCGs were strongly associated with the geographical origin of the *M. tuberculosis* samples and the birthplace of the human hosts. The authors proposed an algorithm able to identify two minimal sets of either 45 or 6 SNPs out of 212 SNPs tested, that could be used for screening of global MTBC collections for studies on evolution, strain differentiation, and biological differences among strains. In another study, Gutacker *et al.,* (2006) studied MTBC genetic relationships by analyzing 36 sSNPs among a big collection of strains from patients enrolled in 4 population-based studies in the United States and Europe, and assigned the strain collection to 1 of 9 major genetic clusters. A similar classification was revealed by analysis of other extended SNPs. Since the classification patterns of the SNP-based phylogenetic lineages were non-randomly associated with IS6110 profiles, spoligotypes, and MIRU-VNTRs, the authors argued for a strongly clonal MTBC population structure.

In parallel, by using DNA microarrays to comprehensively identify large-sequence polymorphisms, a stable association between

Research

MTBC strains and their human host populations was observed (Hirsh *et al.,* 2004); phylogenetic analysis not only indicated that horizontal gene transfers were rare among MTBC, but also that associations between host and pathogen populations were stable even in a cosmopolitan urban setting (like San Francisco), and were largely dictated by the composition of the local immigrant population. The authors concluded that *M. tuberculosis* is organized into several large, genetically differentiated populations, which in turn are directly and stably associated with host populations delineated according to their place of origin. A subsequent report by the same group confirmed this variable host-pathogen compatibility, the global *M. tuberculosis* population structure being defined by six RD/LSP-defined phylogeographical lineages – each associated with specific, sympatric human populations, i.e., the Indo-Oceanic lineage, East-Asian lineage, East-African-Indian lineage, Euro-American lineage, and two West-African lineages (Gagneux *et al.,* 2006.

Table 1. Comparison of spoligotyping-based nomenclature of *M. tuberculosis* lineages vs. PGG groupings, SNPs and SNP-based Cluster groups (SCG), and LSP-based lineages.

Spoligotyping-based (Filliol 2003)	PGG	SCG (Filliol 2006)	SNP-based (Gutacker 2006)	LSP (Gagneux 2006)
East-African-Indian (EAI)	PGG1	SCG 1	sSNP-I	Indo-Oceanic
Beijing	PGG1	SCG 2	sSNP-II	East-Asian
Central-Asian (CAS)	PGG1	SCG 3a	sSNP-IIA	East-African-Indian
Haarlem	PGG2	SCG 3b	sSNP-III	Euro-American
X1	PGG2	SCG 3c	sSNP-IV	Euro-American
X1,X2,X3	PGG2	SCG 4	sSNP-V	Euro-American
LAM	PGG2	SCG 5	sSNP-VI	Euro-American
T (Miscellaneous)	PGG2-3	SCG 6	sSNP-VII sSNP-VIII	Euro-American
Bovis	PGG1	SCG 7	(MTBC)	(MTBC)
M. africanum	PGG1	NA	NA	West-African 1
M. africanum	PGG1	NA	NA	West-African 2

Table 1 summarizes the correspondence among various lineage nomenclatures. It is important to underline that one must keep in mind the marker used when talking about a lineage, particularly for the naming of "East-African Indian" or EAI which denote 2 completely different groups of *M. tuberculosis* by spoligotyping vs. LSPs. Interestingly, a good congruence was observed between spoligotyping and SNPs (Filliol et al., 2006); the East African Indian and Beijing spoligotypes being concordant with SCG-1 and SCG-2, respectively; X and Central Asian spoligotypes were also associated with one SCG or subgroup combination. Other clades had less consistent associations with SCGs. Furthermore, the various spoligotyping-defined lineages fit well with the previous PGG groups, hence MTBC strains can be tentatively classified as ancestral TbD1+/PGG1 group (subset 1: M. africanum and East African Indian, EAI), modern TbD1– /PGG1 group (subset 2: Beijing and Central Asian or CAS), and evolutionary recent TbD1– /PGG2/3 group (subset 3: Haarlem, X, S, T, and Latin American and Mediterranean or LAM). Nonetheless, proper epidemiologic and phylogenetic inferences are not always an easy task due to a lack of understanding of the mechanisms behind the mutations leading to the polymorphism of these genomic targets. Recent studies have shown that phylogenetically unrelated MTBC strains could be sometimes found with the same spoligotype pattern as a result of independent mutational events (Fenner et al., 2011), an observation that corroborates the fact that spoligotyping is prone to homoplasy to a higher extent than the MIRU-VNTRs (Comas et al., 2009). Furthermore, spoligotyping has little discriminative power for families associated with the absence of large blocks of spacers, e.g., the Beijing lineage (Allix-Béguec et al., 2014). For all these reasons, we recommend to make a finer phylogenetic analysis of most significant circulating MTBC clones by multiple genetic markers and compare to the existing data worldwide – a complicated task by itself had it not been for availability of huge international databases that provide with such a framework today.

TB genotyping databases – what is available?

Our knowledge about TB is wider today than ever before, but what is our ability to compare the data generated with respect to all the data that has cumulated over years? Are we really able to instantaneously compare the genetic information on the circulating MTBC strains in conjunction with all demographical, clinical, bacteriological and epidemiological information available in various registers? The necessity of databases in such a context is obvious, and conception and design of databases in the control/surveillance of TB as well as other communicable diseases is certainly going to be an essential tool for achieving the Millennium Development Goals (MDGs) targeted by the WHO for 2015 (WHO, 2006). Indeed, databases allow the storing of huge amount of information in a structured way, facilitating data processing, interrogation, and streamline the decision process thanks to knowledge-based datamining. Nonetheless, they should be constantly updated and maintained like historical heritage and monuments in the present era of « Big Data », representing a more voluminous set of data exceeding the size of traditional databases, a fact which requires for revolutionary measures to be taken for data management, analysis and accessibility of biological data (Howe *et al.,* 2008). In the last couple of years, various databases and web tools have been developed in the TB field mostly devoted to study TB molecular epidemiology and evolution; some examples include:

- SpolDB4 and SITVITWEB are genotyping databases developed at IPG (Brudey *et al.,* 2006; Demay *et al.,* 2012). The later SITVITWEB version is a multimarker database with genotyping data on 62,582 clinical isolates corresponding to 153 countries of patient origin (105 countries of isolation). Method-wise it contains: (a) spoligotyping data, n=7105 patterns from 58180 clinical isolates, grouped in 2740 shared-types or SITs (n=53816 clinical isolates), and 4364 orphan patterns; (b) 12-locus MIRU-VNTRs, n=2379 patterns from 8161 clinical isolates, grouped in 847 shared-types or MITs (n=6626 clinical isolates), and 1533 orphan patterns; (c) 5-locus Exact Tandem Repeats (ETRs), n=458 patterns from 4626 clinical isolates, grouped in 245 shared-types or VITs (n=4413 clinical isolates), and 213 orphan patterns. The SITVITWEB database is freely available at: http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE
- SpolTools is a collection of online browser programs and visualization tool designed to manipulate and analyze MTBC spoligotyping data (Reyes *et al.,* 2008; Tang *et al.,* 2008). It

Research

also contains an online repository of spoligotyped isolates collected from published literature (currently 30 datasets containing 1179 spoligotype patterns corresponding to 6278 isolates). In particular, it allows to draw Spoligoforest trees that illustrate evolutionary relationships between spoligotypes in a given setting. SpolTools is available at: http://www.emi.unsw.edu.au/spolTools/

- MIRU-VNTRplus is a web based tool dedicated to analyze molecular typing data on TB, particularly the 12-, 15- and 24-loci MIRU-VNTRs formats (Allix-Béguec *et al.,* 2008; Weniger *et al.,* 2010). Tools for data exploration include search for similar strains, creation of phylogenetic and minimum spanning trees and mapping of geographic information. In addition, the database also provides detailed results (geographical origin, drug susceptibility profiles, genetic lineages and the spoligotyping pattern, SNP and LSP profiles, and the IS6110-RFLP fingerprints) on a collection of 186 well-characterized reference strains. MIRU-VNTRplus is available at: http://www.miru-vntrplus.org/
- TB Genotyping Information Management System (TB GIMS) is a secure web-based system designed to improve access and dissemination of genotyping information nationwide in the United States (44) [MMWR, 2010]. It stores and manages genotyping data on TB patients in the United States; allows authorized users to submit and track MTBC isolates to and from the contract genotyping labs; provides immediate notification of genotyping results and updates to TB labs and programs; links isolate data to patient-level surveillance data; provides reports on genotype clusters, including national genotype distribution; and provides national, state, and county maps of genotype clusters. This database is not publicly available.
- Mbovis.org is a spoligotype database with over 1400 patterns belonging to following RD9-deleted MTBC lineages: *M. africanum, M. bovis* (antelope), *M. microti, M. pinnipedii, M. caprae* and *M. bovis* (Smith and Upton, 2012). This database is available at: http://www.mbovis.org/
- MycoDB.es is a Spanish database of Animal tuberculosis (Rodriguez-Campos *et al.,* 2012), which was created as an epidemiological tool at national level (Spain). It contains 401 different spoligotype patterns containing 17,273 isolates belonging to *M. bovis, M. caprae* and *M. tuberculosis*, as well as a limited amount of MIRU-VNTR data. Unfortunately, this database is restricted to authorized access, limited to Spanish animal health agency – Centro de Vigilencia Sanitaria Veterinaria (VISAVET): http://www.vigilanciasanitaria.es/mycodb/
- TB-Lineage is an online tool for classification and analysis of MTBC genotypes into major lineages using spoligotypes and optionally MIRU locus 24 (Shabbeer *et al.,* 2012). It was developed and tested using genotyping data from the Centers for Disease Control and Prevention (CDC), Atlanta on 37066 clinical isolates corresponding to 3198 spoligotype patterns and 5430 MIRU-VNTR patterns. However, if MIRU locus 24 data is not available, the system utilizes predictions made by a Naïve Bayes classifier based on spoligotype data alone. The accuracy of automated classification using both spoligotypes and MIRU24 is >99%, and using spoligotypes alone is >95%. TB-Lineage is freely available at http://tbinsight.cs.rpi.edu/run_tb_lineage.html. This website also provides a tool to generate spoligoforests in order to visualize the genetic diversity and relatedness of genotypes and their associated lineages.
- tbvar is a searchable database using a systematic computational pipeline which allows to annotate potential functional and/or drug-resistance-associated variants from clinical re-sequencing data of MTBC (Joshi *et al.,* 2013). For this purpose, the authors re-analyzed re-sequencing datasets corresponding to more than 450 MTBC isolates available in public domain so as to create a comprehensive variome map comprising >29 000 single nucleotide variations. This database can be accessed by browsing location of variants (e.g., 1417019, 3037367, 4222628, etc.); genes (e.g., *katG, pncA, gyrA,* etc.); RvID (e.g., Rv1059, Rv1069c, Rv3693, etc.); or genome position range (10000-15000 ; 30000-35000 ; 80000-85000, etc.); and is available at: http://genome.igib.res.in/tbvar/
- InTB is a web-based interface/system for integrated warehousing and analysis of clinical, socio-demographic and molecular typing data on TB (Soares *et al.,* 2013). It allows to insert and download standard genotyping data in conjunction with an extensive array of clinical and socio-demographic variables that are used to characterize the disease. It also allows to classify new isolates into a well-characterized set of isolates based on internal references, multiple types of data plotting and to generate trees for filtered subsets of data combining molecular and clinical/socio-demographic information. Built on open source software, the full source code along with ready to use packages are available at http://www.evocell.org/inTB.

TB genotyping databases developed at Institut Pasteur de la Guadeloupe (IPG)

The first database was initiated more than fifteen years ago at Institut Pasteur de la Guadeloupe (IPG) when an undergraduate trainee named Jérôme Maïsetti took the initiative of entering the available spoligotype patterns from our own Caribbean isolates (n=218 strains) in an excel spreadsheet and pooled them with published data (n=392 isolates) from other countries. Once the patterns were sorted, we realized that one could not only define predominant patterns but also trace the origin of strains and their potential movements. This database of 610 spoligotypes was tentatively named SpolDB1, and led to the first ever description of 69 major spoligotype patterns in order to better understand TB origin and transmission (Sola *et al.,* 1999). Development of SpolDB1 was followed by the launch of SpolDB2 containing data on 3319 isolates (Sola *et al.,* 2001), and SpolDB3 on 13008 isolates grouped into 813 shared-types (containing 11,708 isolates) and 1300 orphan patterns (Filliol *et al.,* 2002; 2003). More recently, development of the fourth MySQL-based version SpolDB4 in 2006, n=39,295 clinical isolates (Brudey *et al.,* 2006), and SITVITWEB in 2012, n=62,582 clinical isolates (Demay *et al.,* 2012), permitted to have a finer phylogeographical snapshot of circulating MTBC genotypic lineages worldwide. The updated version of our database named SITVIT2 (n=111635 isolates) today contains genotyping information on approximately twice more strains than in the previous version, and will be released in 2014. It should be underlined that these two recent versions are MySQL-based multimarker databases with both spoligotyping and MIRU-VNTR data, limited to 12-loci MIRUs in SITVITWEB; and 12-, 15- or 24-loci MIRU-VNTR data in SITVIT2. In these recent versions, a web-based interface allows the user to search for strains through the database by criteria, such as the year,

Research

the isolation country, the country of origin, the investigator's name; as well as additional combined searches in SITVIT2, making it possible to get the genotyping data on selected strains in conjunction with their geographical distribution, as well as available data on drug-resistance, demographic and epidemiologic characteristics.

These successive developments of IPG databases have allowed to considerably improve our knowledge on genotyping and phylogeny/phylogeography of TB worldwide. The various versions of the our databases have led to a significant number of bilateral and multilateral studies as evidenced by the very high numbers of citations the successive databases have received (interrogation made on Google Scholar on March 31st 2014): SpolDB3 (Filliol *et al.,* 2002; 2003), 180 + 220 or 400 citations collectively; SpolDB4 (Brudey *et al.,* 2006), 661 citations; and SITVITWEB (Demay *et al.,* 2012); 67 citations, although made available only a year ago. We are therefore certain that the next version SITVIT2 will also find its place among the research community as a useful tool not only for TB molecular population genetics, historical demography and epidemiological modeling, but also for fundamental genetic analyses.

Brief description of SITVIT2

Future main functionalities of the SITVIT2 website will be improved as compared to the current SITVITWEB version both numerically as well as at the level of interface for future queries, although the lineage designations will be maintained almost unchanged with few exceptions. At the time of this study, SITVIT2 contained a total of 111,635 MTBC clinical isolates from 169 countries of patient origin. For data collection, we either enriched the database with genotyping results obtained at the institut Pasteur of Guadeloupe, or received from various co-investigators and collaborating laboratories, or those retrieved from published studies (Demay *et al.,* 2012). The website was developed using the java server pages (JSP) technology and embedded in a free Apache Tomcat application server (http://tomcat.apache.org), stored at the Institut Pasteur of Guadeloupe. The java technology was implemented as described earlier (Demay *et al.,* 2012). As previously, the description of the genetic characters of the clinical isolates accessible in SITVIT2 relies on a unique key identifier (IsoNumber) which summarizes information on the country of isolation, a laboratory code number, the year of isolation, a code for drug resistance information (0 to 4), and a unique isolate number given by the participating laboratory/hospital. In accordance with ethical guidelines concerning electronic treatment of data, this allows an anonymous number which can only be decoded by the data provider (the microbiological laboratory which supplied the data) to trace back their patient information but not by other users. SITVIT2 performs the automatized labelling system of SpolDB4 and SITVITWEB that attributes to each spoligotype present in 2 or more strains in the database a Spoligotype International Type (SIT) number, and to each MIRU profile present in 2 or more strains a MIRU International Type (MIT) number. The MIT numbers for 12-, 15- or 24- loci MIRU-VNTRs formats are called as 12-MIT, 15-MIT and 24-MIT while those restricted to 5 Exact Tandem Repeats are labeled as VIT (for VNTR International Type). Note that "orphan" designates patterns reported for a single isolate that does not correspond to any of the patterns recorded in the repository of the SITVIT2 database.

Some main website functionalities include various query tools such as spoligotype format conversion (binary to octal and vice-versa), data submission and analysis for spoligotyping and various MIRU formats, as well as criteria search for each marker individually or combined, year and country of isolation/origin, investigator's name, geographical distribution maps, and associated demographic, epidemiologic, and drug resistance information. It is thus possible to make individual queries or grouped queries by entering a duly formatted excel file (model available on website). For all queries, the user might expect a detailed report on markers (SIT and MIT numbers), phylogenetical lineages, and associated information available in the database in an anonymized format. One of the ongoing developments will include the possibility to look for correspondence of nomenclature of MIT patterns in SITVIT2 according to MIRU-VNTRplus nomenclature (note that a comparison of lineages between the 2 databases did not show major differences; results not shown).

Major phylogenetic lineages in SITVIT2

In SITVIT2, strains are classified in major phylogenetic clades assigned according to signatures provided earlier, which includes various MTBC members (AFRI, *M. africanum*; BOV, *M. bovis*; CANETTII, *M. canettii*; MICROTI, *M. microti*; PINI, M. pinnipedii), as well as for *M. tuberculosis* sensu stricto, i.e., the Beijing clade, the Central-Asian (CAS) clade, the East-African-Indian (EAI) clade, the Haarlem/Ural clades, the Latin-American-Mediterranean (LAM) clade, the Cameroon and Turkey lineages, the «Manu» family, the IS6110-low banding X clade, and the ill-defined T clade. Note that some spoligotypes previously classified as H3/H4 sublineages within Haarlem family were recently relabeled "Ural" (Mokrousov 2012); these include patterns belonging to H4 sublineage that were relabeled «Ural-2», and some patterns previously classified as H3 sublineage but with an additional specific signature (presence of spacer 2, absence of spacers 29 to 31, and 33 to 36), that are now relabeled «Ural-1». Furthermore, two LAM sublineages were recently raised to independent lineage level: LAM10-CAM as Cameroon lineage (Koro Koro *et al.,* 2013), and LAM7-TUR as Turkey lineage (Abadia *et al.,* 2010; Kisa *et al.,* 2012). We have kept this nomenclature unaltered for spoligotyping based genotypic lineages since it has already been found to be useful for local or global molecular epidemiological studies, as well as to follow the evolutionary and quantitative genetics of tubercle bacilli at a global scale. Note that the distribution of clinical isolates in SITVIT2 is studied both country wise and at macro-geographical level as sub-regions defined according to United Nations (according to http://unstats.un.org/unsd/methods/m49/m49regin.htm); regions: AFRI (Africa), AMER (Americas), ASIA (Asia), EURO (Europe), and OCE (Oceania), subdivided in: E (Eastern), M (Middle), C (Central), N (Northern), S (Southern), SE (South-Eastern), and W (Western). In this classification scheme, CARIB (Caribbean) belongs to Americas, while Oceania is subdivided in 4 sub-regions, AUST (Australasia), MEL (Melanesia), MIC (Micronesia), and POLY (Polynesia). Note that Russia was attributed a new sub-region by itself (Northern Asia) instead of including it among the rest of Eastern Europe.

The readers are requested to refer to **Table 2** for a brief summary on the comparison of SITVITWEB vs. SITVIT2 and corresponding phylogenetical lineages in the 2 versions, and to **Figure 1** which highlights the evolution of strains recorded by various geographical sub-regions between the 2 versions. The most noticeable increase between the 2 versions can be observed for Southern Europe and Eastern Asia, followed

Research

by Central and Southern America, Northern, Eastern and Western Africa, and Northern and Western Europe (**Figure 1**). Lineage-wise, the proportion of Beijing genotype did not differ significantly between SITVITWEB and SITVIT2 (representing respectively 9.84% vs. 9.72% of isolates globally; **Table 2**); and was predominant in Asia, followed by significantly visible proportions in North America, South Africa, and Australasia. Proportions were also similar for CAS (3.69% vs. 3.91%), Cameroon (previously LAM10-CAM: 1.04% vs. 0.98%), Turkey (previously LAM7-TUR: 0.59% vs. 0.53%), and Manu (1.08% vs. 0.95%) lineages. However the most significant increase was observed for *M. bovis* strains (10.36% vs. 23.06%), underlining the increased potential of SITVIT2 not only to study M. *tuberculosis* epidemiology but also bovine tuberculosis.

Table 2. A summarized representation of the SITVITWEB and SITVIT2 databases and the corresponding major phylogenetic lineages of the M. tuberculosis complex (MTBC).

Major Lineages*	SITVITWEB * (n=62582)		SITVIT2 (n=111635)	
	Nb	%	Nb	%
Beijing	6,159	9.84	10,850	9.72
AFRI	695	1.11	965	0.86
BOV	6,486	10.36	25,741	23.06
CANETTII	12	0.02	12	0.01
CAS	2,480	3.96	4,362	3.91
EAI	4,674	7.47	6,617	5.93
Haarlem/Ural	7,058	11.28	10,580	9.48
LAM	8,042	12.85	12,245	10.97
Cameroon (previously LAM10-CAM)	650	1.04	1095	0.98
Turkey (previously LAM7-TUR)	370	0.59	593	0.53
Manu	675	1.08	1,064	0.95
MICROTI	29	0.05	29	0.03
PINI	152	0.24	159	0.14
S	1,151	1.84	1,606	1.44
T	12,038	19.24	17,947	16.08
X	4,088	6.53	4,683	4.19

Worldwide distribution maps in SITVIT2

The worldwide distribution map of major lineages in SITVIT2 illustrated in **Figure 2** highlights the global phylogeographical geo-specificities of MTBC isolates as seen in a 2014 snapshot. Even though these geo-specificities were already suggested in previous studies based both on spoligotyping (Brudey *et al.,* 2006; Demay *et al.,* 2012); and LSPs (Gagneux *et al.,* 2006), the present map corroborates the current distribution pattern and specificities thanks to curated data in SITVIT2 database. Furthermore, the reclassification of certain lineages that didn't appear in previous versions such as "Ural" shows that it is noticeably present in Russia, Central Asia, Southern Asia, and Western Asia, as well as in Finland (Northern Europe) which shares a common frontier and privileged links to Russia, specially as a buffer zone in a succession of wars between Russia and Sweden during the 18th century (http://www.historyworld.net/wrldhis/PlainTextHistories.asp?historyid=ad02). We further confirmed the relabeling of LAM7-TUR as Turkey lineage and that of LAM10-CAM as Cameroon lineage thanks to their phylogeographical distribution in SITVIT2 (**Figure 2**). It is important to underline that the Turkey lineage is progressing in Eastern European countries (representing around 6% of all MTBC strains in Eastern Europe in 2014 vs. less than 2% until 2006; results not shown). One may also notice the continuous reduction of the proportion of AFRI lineage in Western Africa, which represented around 37% of all MTBC strains in this sub-region until 2006 vs. 29% in 2014. This observation corroborates our previous suggestion that evolutionary ancestral *M. africanum* strains in Western Africa are slowly being replaced by evolutionary recent MTBC lineages such as Cameroon or other Euro-American lineages (Groenheit *et al.,* 2011). Last but not least, although the global proportions of CAS did not significantly change between SITVITWEB and SITVIT2 (3.69% vs. 3.91%), an increase was observed in Western Asia (12% vs. 18%) and East Africa (10% vs. 15%).

Looking for major associations between phylogenetic lineages vs. demographic and epidemiologic characteristics in SITVIT2

A user will be able to retrieve phylogenetic information linked to a number of different parameters such as incidence of the disease as seen through maps available from WHO, demography (age, sex-ratio) and other characteristics from various sub-regions in order to underline specificities in function of countries, regions, and populations. Considering that spreading of MDR- and XDR-TB clones in general populations represents a major threat in TB control, the interest of such a database is also to provide with a prevailing snapshot as well as to pinpoint any emergence for Public Health authorities. This geo-referencing of our data using Google API already provides with a potential Geographical Information Systems (GIS), that represents an efficient bacteriological counterpart of the WHO's Communicable Disease Global Atlas (http://apps.who.int/globalatlas/), allowing to bring together in a single electronic platform analysis and interpretation of genotyping data in conjunction with information on demography, socioeconomic conditions, and environmental factors. In our opinion, one should use such a mapping ideally in conjunction with suitable statistical and bioinformatical tools and softwares to better describe the TB genetic landscape. In addition to the tools and softwares described in previous studies (Brudey *et al.,* 2006; Demay *et al.,* 2012); we also use following tools in SITVIT2 analyses performed in routine:

(a) STATA software version 12 for descriptive and univariate analyses.

(b) R software version 2.14.1 to calculate the Odds Ratios (OR) and 95% Confidence Interval (CI) values.

(c) Pearson's Chi-square test and Fisher's Exact Test to compare major associations between genotyping data (shared types /

*The strains are classified in major phylogenetic clades assigned according to signatures provided earlier (Demay et al., 2012); which includes various MTBC members (AFRI, *M. africanum*; BOV, *M. bovis*; CANETTII, *M. canettii;* MICROTI, *M. microti*; PINI, *M. pinnipedii*), as well as for lineages/sub-lineages of *M. tuberculosis* sensu stricto (note that the sublineages are not shown), i.e., the Beijing clade, the Central-Asian (CAS) clade, the East-African-Indian (EAI) clade, the Haarlem/Ural clades,the Latin-American-Mediterranean (LAM) clade, the Cameroon and Turkey lineages, the «Manu» family, the IS6110-low banding X clade, and the ill-defined T clade.

Research

lineages) vs. demographic, epidemiologic, or socioeconomic characteristics (P values of <0.05 being considered as statistically significant).
(d) Minimum Spanning Trees (MSTs) are constructed based on genotyping data (spoligotypes, MIRUs) using BioNumerics software, version 6.6 (Applied Maths, Sint-Martens-Latem, Belgium). MSTs are connected undirected graphs in which all of the patterns are linked together with the fewest possible linkages between nearest neighbors.
(e) SpolTools software (http://www.emi.unsw.edu.au/spolTools) is used to draw Spoligoforests trees based on the Fruchterman-Reingold algorithm or a Hierarchical Layout (Reyes *et al.,* 2008; Tang *et al.,* 2008). Note that contrary to the MSTs, Spoligoforests are directed (and not necessarily connected) graphs allowing to highlight the evolutionary relationships between ascendant and descendant spoligotyping patterns.
(f) GraphViz software available at http://www.graphviz.org (Ellson *et al.,* 2002) to color the Spoligoforests in function of the lineages.
(g) WebLogo application version 2.8.2 (available at http://weblogo.berkeley.edu/(Schneider and Stephens, 1990; Crooks *et al.,* 2004) to evaluate and visualize the allelic diversity of the spoligotyping patterns in function of their associated lineages. This method of representation adapted to 43-spacer spoligotyping was labeled as "Spoligologos" (Driscoll *et al.,* 2002). WebLogos consist of stacks of symbols as graphical representations of the comparative diversity observed for individual spoligotyping spacer – one stack for each of the 43 spacer. The overall height of the stack indicates the conservation of a given spacer (the letter "n" designates the presence of a spacer, and letter "o" designates the absence). If a spacer is always the present or absent for one position of the 43 available ones (i.e., if 100% of the strains conserve the same presence or absence in one position, it corresponds to 4 bits), while the height of individual symbols within the stack indicates the relative frequency of absent/present spacer at that position.

Examples of some recent studies using SITVIT2 database

It is obviously not possible to detail all the studies done using SITVIT2, however some recent examples that looked for associations between phylogenetic lineages vs. demographic and epidemiologic parameters include:
(a) Geographical distribution map of spoligotyping-based MTBC lineages in various subregions of Africa and high phylogeographical specificity of *M. africanum* for Western Africa, with Guinea-Bissau being the epicenter (Groenheit *et al.,* 2011).
(b) Evidence that MTBC strains potentially involved in the TB epidemic in Sweden a century ago belonged to a closely knit pool of evolutionary recent PGG2/3 strains restricted to Sweden and its immediate neighbors (Groenheit *et al.,* 2012).
(c) Exploration of MTBC phylogenetic associations with drug resistance in Peru suggesting a prolonged, clonal, hospital-based outbreak of MDR disease amongst HIV patients (Sheen *et al.,* 2013).

(d) Phylogeographical mapping of TB in Finland showing a close resemblance of global MTBC population structure to one reported for Sweden, specially the predominance of the Euro-American family among elderly TB patients; the main difference being observed for the Ural lineage which was present in significant proportions among Finnish born cases (and also found in Russia, Latvia, and Estonia), but not in Sweden (Smit *et al.,* 2013).
(e) At the worldwide level using the SITVIT2 database, we find that X and LAM lineages are significantly more associated with HIV-positive serology; p-value<0.0001 (article in preparation).
(f) In several studies, the association of Beijing lineage with excessive drug resistance including MDR and/or XDR-TB was highlighted (van Soolingen *et al.,* 1995; Glynn *et al.,* 2002; Parwati *et al.,* 2010). Hence we recently performed exploration of phylogenetic associations (MTBC split in 2 groups as Beijing vs. other lineages) with drug resistance (quantified as pansusceptible, MDR-TB, XDR-TB, or any other drug resistance) using the SITVIT2 database (Couvin and Rastogi, 2014). The distribution of drug resistance for different subregions is shown in **Figure 3A**. Although proportion of drug-resistant strains was significantly higher for Beijing vs. non Beijing strains globally, important variations in the distribution of drug-resistance were observed. Drug resistance was significantly linked to Beijing (vs. non-Beijing strains) in Russia, Southern Asia, Southeastern Asia, and European countries, but not in Americas, Western Asia, China and Japan. If one considers the evolution of drug resistance over time for Beijing strains (1998 to 2011, **Figure 3B**), a continuous progression in the proportion of MDR and XDR strains (and a relative decrease of pansusceptible strains) is visible worldwide since 2003. Last but not least, we also observed that a rare but emerging spoligotype pattern (SIT190/Beijing) was significantly more associated with MDR-TB than the traditional SIT1/Beijing pattern (p-value<0.0001).
(g) Regarding the information provided thanks to a Spoligoforest tree, we would refer to a recent study conducted in Baghdad, Iraq (Mustafa Ali *et al.,* 2014). The results obtained on a total of 270 MTBC isolates showed that 2 specific patterns SIT1144/T1 and SIT309/CAS1-Delhi predominated in this study (6.3% for each pattern). The evolutionary relationships between Iraqi isolates as seen through a Spoligoforest tree using a Hierarchical Layout (**Figure 4**) clearly show that the bulk of TB in postwar Iraq is limited to 2 phylogenetically related group of MTBC strains belonging to T and CAS lineages.
(h) Regarding the information provided thanks to geo-referencing of genotyping data in SITVIT2 using Google API, we could refer to the same study (Mustafa Ali *et al.,* 2014). As shown in **Figure 5**, significant differences were highlighted between Baghdad city as compared to other cities in Iraq, regarding both demographics and drug resistance information. Indeed, with a male/female sex ratio of 1.49 in Baghdad vs. 3.04 in other governorates, the proportion of female patients was significantly higher in Baghdad city (p-value = 0.009; Odds Ratio = 0.49 and 95%CI [0.28; 0.86]). The proportion of newly treated vs. re-treated cases differed significantly between the 2 groups, the proportion of re-treated patients being higher in other governorates of Iraq (p-value=0.007; OR=2.1, 95%CI [1.19; 3.62]). Finally, the rate of MDR-TB was higher in other governorates than Baghdad (p >0.12; difference not statistically significant).

Conclusions

The collection of databases developed at Institut Pasteur de la

Research

Figure 1. (A) Worldwide distribution of MTBC isolates recorded in both SITVITWEB (blue circle) and SITVIT2 (green circle), the number of isolates per sub-region is indicated inside each circle. (B). Gradual evolution of series of databases (from SpolDB2 to SITVIT2) in function of the number of MTBC isolates for each database.

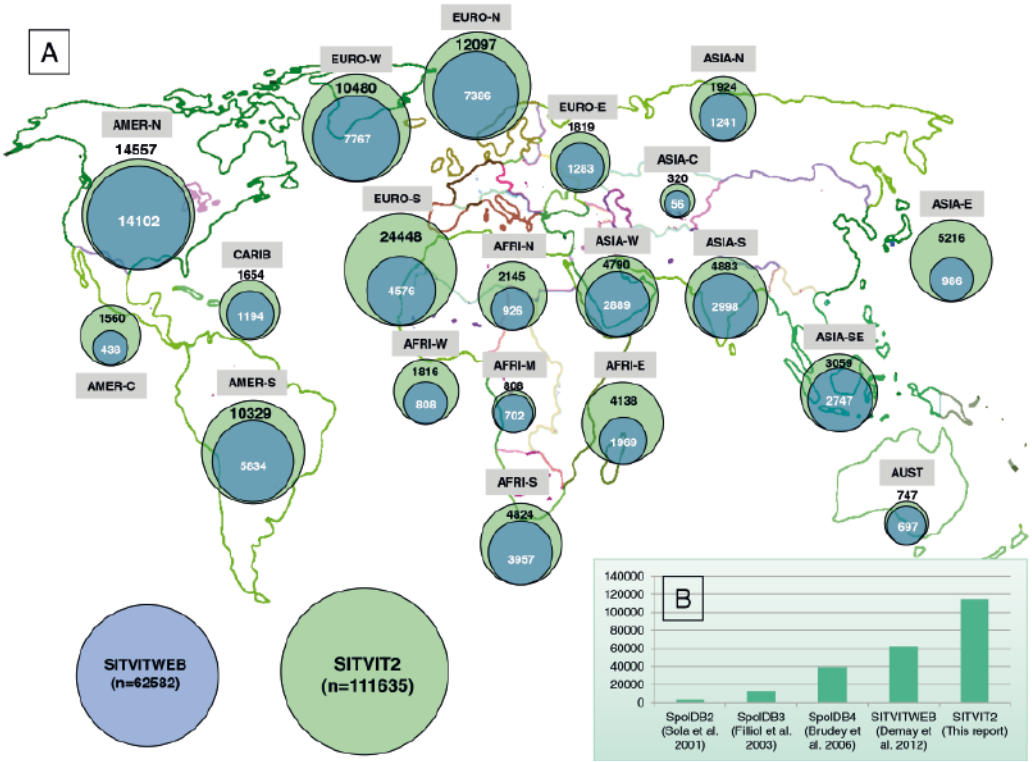
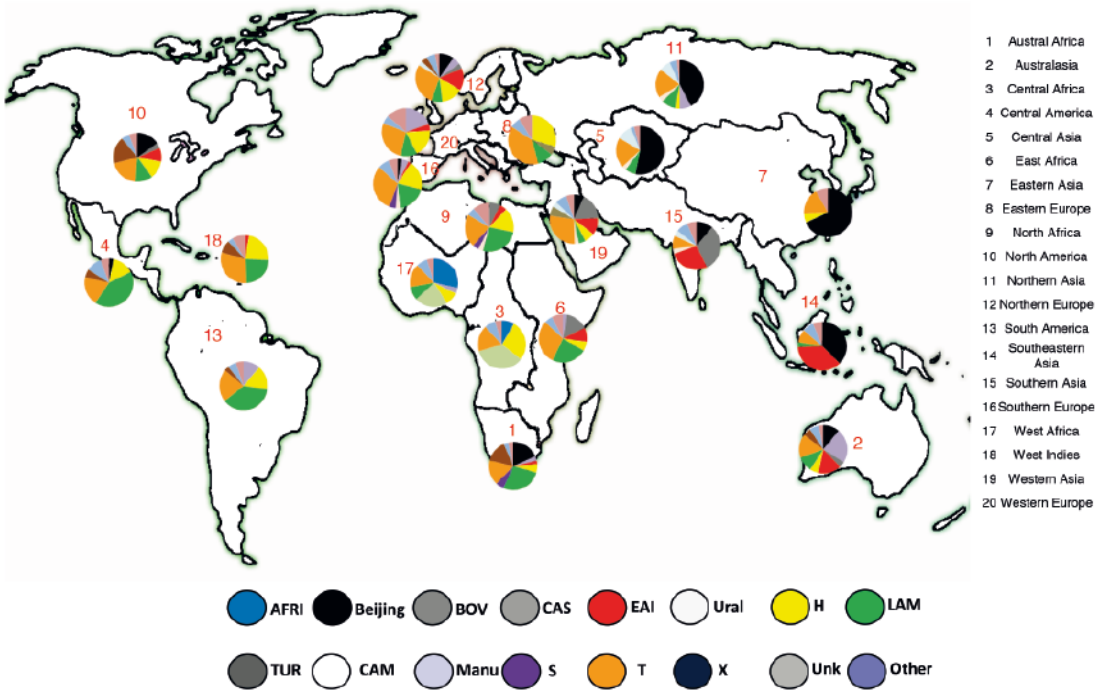


Figure 2 : Worldwide distribution of lineages contained in the SITVIT2 database.



Research

Guadeloupe has helped achieve a global overview of worldwide TB situation. With the newly updated SITVIT2 database, we have been further able to highlight the ongoing circulation of MTBC strains based on extended genotyping markers, as well as underlined major associations between MTBC phylogenetic lineages vs. demographic and epidemiologic characteristics. Future developments should ideally include inclusion of other markers such as RD/LSP and SNPs, as well as the future information that is going to be generated thanks to next generation sequencing. Although many evolutionary and pathobiological characteristics of the prevailing TB epidemic remain to be discovered, the new SITVIT collection of databases represents a major tool for an improved epidemiological surveillance and control of TB.

Acknowledgements

We thank more than 500 investigators who provided data for SpolDB4, SITVITWEB, and SITVIT2 databases (among which 190 contributed genotyping data on more than 100 strains). Note that each strain with its genotyping information in these respective databases directly refers to the investigator in question (the full list is available upon request). We are highly grateful to Thierry Zozio, Julie Millet, Veronique Hill, and Elisabeth Streit for helpful discussions. DC was awarded a Ph.D. fellowship by the European Social Funds through the Regional Council of Guadeloupe.

References

Abadia E, Zhang J, dos Vultos T, Ritacco V, Kremer K, Aktas E, Matsumoto T, Refregier G, van Soolingen D, Gicquel B, Sola C. (2010). Resolving lineage assignation on *Mycobacterium tuberculosis* clinical isolates classified by spoligotyping with a new high-throughput 3R SNPs based method. *Infect Genet Evol*. 2010; 107: 1066–1074.

Allix-Béguec C, Harmsen D, Weniger T, Supply P, Niemann S. 2008. Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. *J Clin Microbiol*. 46: 2692-2699.

Allix-Béguec C, Wahl C, Hanekom M, Nikolayevskyy V, Drobniewski F, Maeda S, Campos-Herrero I, Mokrousov I, Niemann S, Kontsevaya I, Rastogi N, Samper S, Sng LH, Warren RM, Supply P. 2014. Proposal of a consensus set of hypervariable mycobacterial interspersed repetitive-unit-variable-number tandem-repeat loci for subtyping of *Mycobacterium tuberculosis* Beijing isolates. *J Clin Microbiol*. 52: 164-172.

Barry CE 3rd, Boshoff HI, Dartois V, Dick T, Ehrt S, Flynn J, Schnappinger D, Wilkinson RJ, Young D. 2009. The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat Rev Microbiol*. 7: 845-855.

de Boer AS, Borgdorff MW, de Haas PE, Nagelkerke NJ, van Embden JD, van Soolingen D. 1999. Analysis of rate of change of IS6110 RFLP patterns of *Mycobacterium tuberculosis* based on serial patient isolates. *J Infect Dis*. 180: 1238-1244.

Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K, Garnier T, Gutierrez C, Hewinson G, Kremer K, Parsons LM, Pym AS, Samper S, van Soolingen D, Cole ST. 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci USA*. 99: 3684-3689.

Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajoj SA, Allix C, Aristimuño L, Arora J, Baumanis V, Binder L, Cafrune P, Cataldi A, Cheong S, Diel R, Ellermeier C, Evans JT, Fauville-Dufaux M, Ferdinand S, Garcia de Viedma D, Garzelli C, Gazzola L, Gomes HM, Guttierez MC, Hawkey PM, van Helden PD, Kadival GV, Kreiswirth BN, Kremer K, Kubin M, Kulkarni SP, Liens B, Lillebaek T, Ho ML, Martin C, Martin C, Mokrousov I, Narvskaïa O, Ngeow YF, Naumann L, Niemann S, Parwati I, Rahim Z, Rasolofo-Razanamparany V, Rasolonavalona T, Rossetti ML, Rüsch-Gerdes S, Sajduda A, Samper S, Shemyakin IG, Singh UB, Somoskovi A, Skuce RA, van Soolingen D, Streicher EM, Suffys PN, Tortoli E, Tracevska T, Vincent V, Victor TC, Warren RM, Yap SF, Zaman K, Portaels F, Rastogi N, Sola C. 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC Microbiol. 6: 23.

CDC. 2010. Launch of TB Genotyping Information Management System (TB GIMS). *Morbidity and Mortality Weekly Report* (MMWR) March 19, 59(10); 300.

Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, Parkhill J, Malla B, Berg S, Thwaites G, Yeboah-Manu D, Bothamley G, Mei J, Wei L, Bentley S, Harris SR, Niemann S, Diel R, Aseffa A, Gao Q, Young D, Gagneux S. 2013. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat Genet*. 45: 1176-1182.

Comas I, Homolka S, Niemann S, Gagneux S. 2009. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS One*. 4: e7815.

Couvin D, Rastogi N. (2014). Tuberculosis – a global emergency: tools and methods to monitor, understand, and control the epidemic with specific example of the Beijing lineage. Tuberculosis (Edinb). Submitted for publication.

Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. *Genome Res*. 14: 1188-1190.

Dale JW, Brittain D, Cataldi AA, Cousins D, Crawford JT, Driscoll J, Heersma H, Lillebaek T, Quitugua T, Rastogi N, Skuce RA, Sola C, Van Soolingen D, Vincent V. 2001. Spacer oligonucleotide typing of bacteria of the *Mycobacterium tuberculosis* complex: recommendations for standardised nomenclature. *Int J Tuberc Lung* Dis. 5: 216-219.

Demay C, Liens B, Burguière T, Hill V, Couvin D, Millet J, Mokrousov I, Sola C, Zozio T, Rastogi N. 2012. SITVITWEB – a publicly available international multimarker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. *Infect Genet Evol*. 12: 755-766.

Driscoll JR, Bifani PJ, Mathema B, McGarry MA, Zickas GM, Kreiswirth BN, Taber HW. 2002. Spoligologos: a bioinformatic approach to displaying and analyzing *Mycobacterium tuberculosis* data. *Emerg Infect Dis*. 8(11): 1306-9.

Ellson J, Gansner E, Koutsofios L, North SC, Woodhull G. 2002. Graphviz – Open Source Graph Drawing Tools. In: Mutzel P, Jünger M, Leipert S (Editors), Heidelberg: Springer-Verlag Berlin. pp. 483-484.

van Embden JDA, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, Hermans P, Martin C, McAdam R, Shinnick TM, Small PM. 1993. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol*. 31, 406-409.

Fang Z, Kenna DT, Doig C, Smittipat DN, Palittapongarnpim P, Watt B, Forbes KJ. 2001. Molecular evidence for independent occurrence of IS6110 insertions at the same sites of the genome of *Mycobacterium tuberculosis* in different clinical isolates. *J Bacteriol*. 183: 5279-5284.

Fenner L, Malla B, Ninet B, Dubuis O, Stucki D, Borrell S, Huna T, Bodmer T, Egger M, Gagneux S. (2011) "Pseudo-Beijing": evidence for convergent evolution in the direct repeat region of *Mycobacterium tuberculosis*. *PLoS One*. 6: e24737.

Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbón MH, Bobadilla del Valle M, Fyfe J, García-García L, Rastogi N, Sola C, Zozio T, Guerrero MI, León CI, Crabtree J, Angiuoli S, Eisenach KD, Durmaz R, Joloba ML, Rendón A, Sifuentes-Osornio J, Ponce de León A, Cave MD, Fleischmann R, Whittam TS, Alland D. 2006. Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for *a minimal* standard SNP set. *J Bacteriol*. 2006; 188: 759-772.

Filliol I, Driscoll JR, Van Soolingen D, Kreiswirth BN, Kremer K, Valétudie G, Anh DD, Barlow R, Banerjee D, Bifani PJ, Brudey K, Cataldi A, Cooksey RC, Cousins DV, Dale JW, Dellagostin OA, Drobniewski F, Engelmann G, Ferdinand S, Gascoyne-Binzi D, Gordon M, Gutierrez

Research

Figure 3. Drug resistance characteristics of Beijing vs. non-Beijing *M. tuberculosis* lineages (A) and evolution of drug resistance among Beijing isolates between 1998 to 2011 (B).

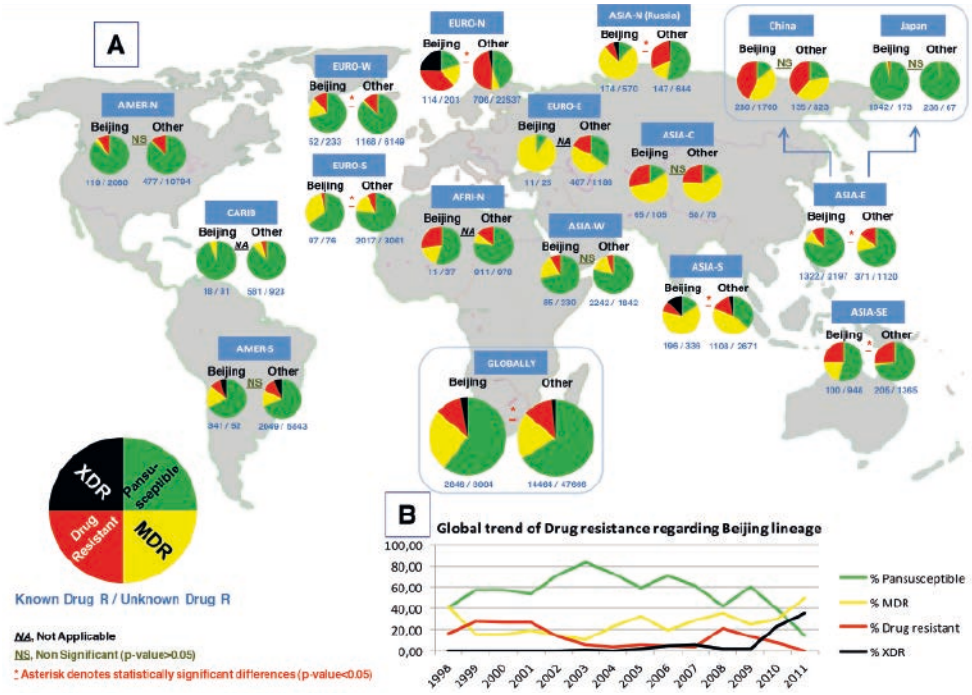
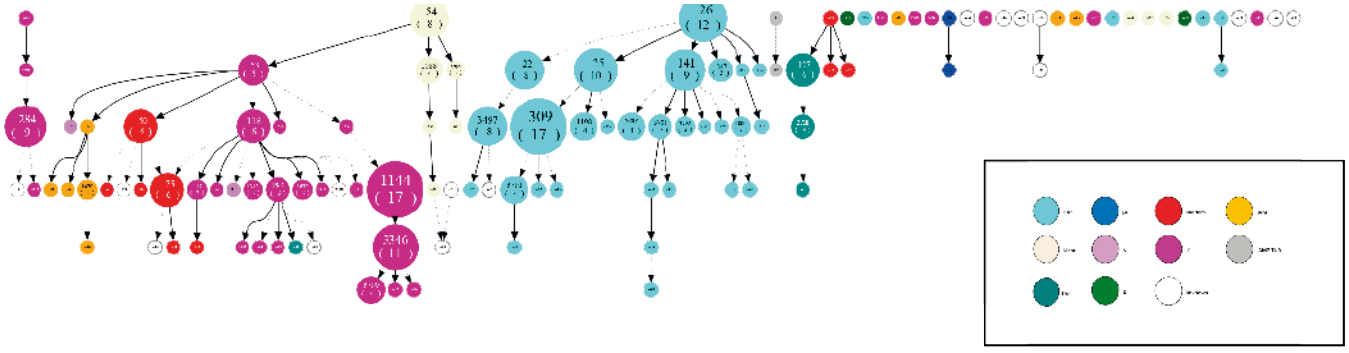
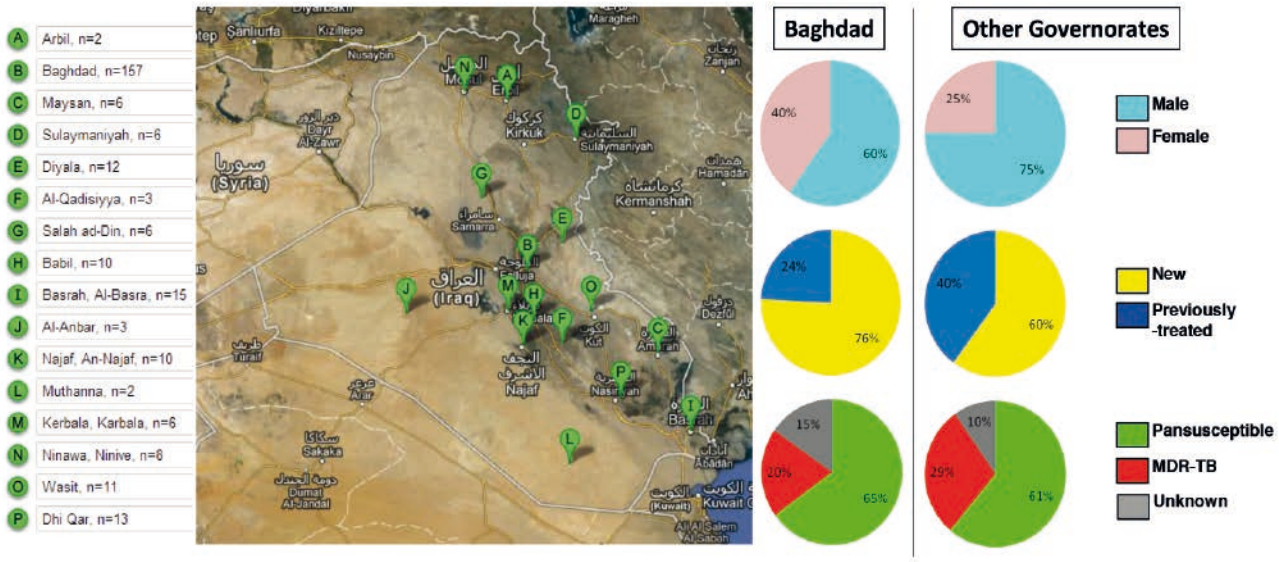


Figure 4. A Spoligoforest representation of parent to descendant spoligotypes in a study from Baghdad, Iraq (n=270 isolates) using a Hierarchical Layout. In this tree, each spoligotype pattern is represented by a node with area size being proportional to the total number of isolates with that specific pattern. Changes (loss of spacers) are represented by directed edges between nodes, with the arrowheads pointing to descendant spoligotypes. The heuristic used selects a single inbound edge with a maximum weight using a Zipf model. Solid black lines link patterns that are very similar, i.e., loss of one spacer only (maximum weigh being 1.0), while dashed lines represent links of weight comprised between 0.5 and 1, and dotted lines a weight less than 0.5. Note that SIT309/CAS1-Delhi and SIT1144/T1 are the biggest nodes (n=17), followed by SIT26/CAS1-Delhi (n=12), SIT3346/T1 (n=11) and SIT25/CAS1-Delhi (n=10), which are other predominant patterns in our study. Finally, orphan isolates (double circled), appear mostly at terminal positions on the tree, or are isolated strains without interconnections with the other strains (figure based on data from reference 70).



Research

Figure 5. A Google Map screenshot showing MTBC strain distribution in function of city of isolation in Baghdad vs. other cities in Iraq (n=270 isolates) and comparison of data in function of patient origin vs. sex-ratio, treatment status, and drug resistance for Baghdad vs. other governorates (figure based on data from reference 70).



MC, Haas WH, Heersma H, Källenius G, Kassa-Kelembho E, Koivula T, Ly HM, Makristathis A, Mammina C, Martin G, Moström P, Mokrousov I, Narbonne V, Narvskaya O, Nastasi A, Niobe-Eyangoh SN, Pape JW, Rasolofo-Razanamparany V, Ridell M, Rossetti ML, Stauffer F, Suffys PN, Takiff H, Texier-Maugein J, Vincent V, De Waard JH, Sola C, Rastogi N. 2002. Global distribution of *Mycobacterium tuberculosis* spoligotypes. *Emerg Infect Dis*. 2002 Nov; 8(11): 1347-9.

Filliol I, Driscoll JR, van Soolingen D, Kreiswirth BN, Kremer K, Valétudie G, Dang DA, Barlow R, Banerjee D, Bifani PJ, Brudey K, Cataldi A, Cooksey RC, Cousins DV, Dale JW, Dellagostin OA, Drobniewski F, Engelmann G, Ferdinand S, Gascoyne-Binzi D, Gordon M, Gutierrez MC, Haas WH, Heersma H, Kassa-Kelembho E, Ho ML, Makristathis A, Mammina C, Martin G, Moström P, Mokrousov I, Narbonne V, Narvskaya O, Nastasi A, Niobe-Eyangoh SN, Pape JW, Rasolofo-Razanamparany V, Ridell M, Rossetti ML, Stauffer F, Suffys PN, Takiff H, Texier-Maugein J, Vincent V, de Waard JH, Sola C, Rastogi N. 2003. Snapshot of moving and expanding clones of *Mycobacterium tuberculosis* and their global distribution assessed by spoligotyping in an international study. *J Clin Microbiol*. 41(5): 1963-1970.

Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S, Nicol M, Niemann S, Kremer K, Gutierrez MC, Hilty M, Hopewell PC, Small PM. 2006. Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA*. 103(8): 2869-73.

García de Viedma D, Mokrousov I, Rastogi N. 2011. Innovations in the molecular epidemiology of tuberculosis. *Enferm Infecc Microbiol* Clin. 29 (Suppl 1): 8-13.

Glynn JR, Whiteley J, Bifani PJ, Kremer K, van Soolingen D. (2002). Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. *Emerg Infect Dis*. 8(8): 843-849.

Gordon SV, Brosch R, Billault A, Garnier T, Eiglmeier K, Cole ST. 1999. Identification of variable regions in the genomes of tubercle bacilli using bacterial artificial chromosome arrays. *Mol Microbiol*. 32: 643-655.

Groenheit R, Ghebremichael S, Svensson J, Rabna P, Colombatti R, Riccardi F, Couvin D, Hill V, Rastogi N, Koivula T, Källenius G. 2011. The Guinea-Bissau family of *Mycobacterium tuberculosis* complex revisited. *PLoS One*. 6(4): e18601.

Groenheit R, Ghebremichael S, Pennhag A, Jonsson J, Hoffner S, Couvin D, Koivula T, Rastogi N, Källenius G. 2012. *Mycobacterium tuberculosis* strains potentially involved in the TB epidemic in Sweden a century ago. *PLoS One*. 7(10): e46848.

Gutacker MM, Mathema B, Soini H, Shashkina E, Kreiswirth BN, Graviss EA, Musser JM. Single-nucleotide polymorphism-based population genetic analysis of *Mycobacterium tuberculosis* strains from 4 geographic sites. *J Infect Dis*. 2006; 193: 121-128.

Heersma HF, Kremer K, van Embden JD. 1998. Computer analysis of IS6110 RFLP patterns of *Mycobacterium tuberculosis*. *Methods Mol Biol*. 101: 395-422.

Hermans PW, van Soolingen D, Bik EM, de Haas PE, Dale JW, van Embden JD. 1991. Insertion element IS987 from *Mycobacterium bovis* BCG is located in a hot-spot integration region for insertion elements in *Mycobacterium tuberculosis* complex strains. *Infect Immun*. 59: 2695-2705.

Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, Homolka S, Roach JC, Kremer K, Petrov DA, Feldman MW, Gagneux S. 2008. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol*. 2008, 6(12): e311.

Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small PM. Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. *Proc Natl Acad Sci USA*. 2004; 101: 4871-4876.

Howe D, Costanzo M, Fey P, Gojobori T, Hannick L, Hide W, Hill DP, Kania R, Schaeffer M, St Pierre S, Twigger S, White O, Rhee SY. (2008) Big data: The future of biocuration. *Nature*. 455(7209): 47-50.

Jagielski T, van Ingen J, Rastogi N, Dziadek J, Mazur PK, Bielecki J. 2014. Current Methods in the Molecular Typing of *Mycobacterium tuberculosis* and Other Mycobacteria. *Biomed Res Int*. 2014 (Article ID 645802): 1-21.

Joshi KR, Dhiman H, Scaria V. (2013) tbvar: a comprehensive genome variation resource for *Mycobacterium tuberculosis*. Database Vol. 2013: article ID bat083.

Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, van

Research

Embden J. 1997. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol*. 35: 907-914.

Kato-Maeda M, Rhee JT, Gingeras TR, Salamon H, Drenkow J, Smittipat N, Small PM. 2001. Comparing genomes within the species *Mycobacterium tuberculosis*. *Genome Res*. 11: 547-554.

Kisa O, Tarhan G, Gunal S, Albay A, Durmaz R, Saribas Z, Zozio T, Alp A, Ceyhan I, Tombak A, Rastogi N. (2012). Distribution of spoligotyping defined genotypic lineages among drug-resistant *Mycobacterium tuberculosis* complex clinical isolates in Ankara, Turkey. *PLoS One*. 2012; 7(1): e30331.

Koro Koro F, Kamdem Simo Y, Piam FF, Noeske J, Gutierrez C, Kuaban C, Eyangoh SI. 2013. Population dynamics of *tuberculous Bacilli* in Cameroon as assessed by spoligotyping. *J Clin Microbiol.* 51: 299-302.

Mahairas GG, Sabo PJ, Hickey MJ, Singh DC, Stover CK. 1996. Molecular analysis of genetic differences between *Mycobacterium bovis* BCG and virulent M. bovis. J Bacteriol. 178: 1274-1282.

Mokrousov I. 2012. The quiet and controversial: Ural family of *Mycobacterium tuberculosis*. *Infect Genet Evol*. 12: 619-629.

Mustafa Ali R, Trovato A, Couvin D, Al-Thwani AN, Borroni E, Dhaer FH, Rastogi N, Cirillo DM. 2014. Molecular Epidemiology and Genotyping of *Mycobacterium tuberculosis* Isolated in Baghdad. BioMed Research International, vol. 2014, Article ID 580981, 15 pages.

Parwati I, van Crevel R, van Soolingen D. 2010. Possible underlying mechanisms for successful emergence of the *Mycobacterium tuberculosis* Beijing genotype strains. *Lancet Infect Dis*. 10: 103-111.

Radhakrishnan I, K MY, Kumar RA, Mundayoor S. 2001. Implications of low frequency of IS6110 in fingerprinting field isolates of *Mycobacterium tuberculosis* from Kerala, India. *J Clin Microbiol*. 39: 1683.

Rastogi N, Sola C. 2007. Chapter 2 – Molecular evolution of the *Mycobacterium tuberculosis* complex. In Tuberculosis 2007: from basic science to patient care, Edited by Palomino JC, Leao S, Ritacco V. 2007, 53-91, Amedeo Online Textbooks; http://pdf.flyingpublisher.com/tuberculosis2007.pdf

Reyes JF, Francis AR, Tanaka MM. 2008. Models of deletion for visualizing bacterial variation: an application to tuberculosis spoligotypes. *BMC Bioinformatics*. 9: 496.

Rodriguez-Campos S, González S, de Juan L, Romero B, Bezos J, Casal C, Álvarez J, Fernández-de-Mera IG, Castellanos E, Mateos A, Sáez-Llorente JL, Domínguez L, Aranaz A; Spanish Network on Surveillance Monitoring of Animal Tuberculosis. 2012. A database for animal tuberculosis (mycoDB.es) within the context of the Spanish national program for eradication of bovine tuberculosis. *Infect Genet Evol.* 12(4): 877-82.

Rothschild BM, Martin LD, Lev G, Bercovier H, Bar-Gal GK, Greenblatt C, Donoghue H, Spigelman M, Brittain D. 2001. *Mycobacterium tuberculosis* complex DNA from an extinct bison dated 17,000 years before the present. *Clin Infect Dis*. 33: 305-311.

Salamon H, Segal MR, Ponce de Leon A, Small PM. 1998. Accommodating error analysis in comparison and clustering of molecular fingerprints. *Emerg Infect Dis*. 4: 159-168.

Schneider TD, Stephens RM. 1990. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. 18: 6097-6100.

Shabbeer A, Cowan LS, Ozcaglar C, Rastogi N, Vandenberg SL, Yener B, Bennett KP. 2012. TB-Lineage: an online tool for classification and analysis of strains of *Mycobacterium tuberculosis* complex. *Infect Genet Evol.* 12: 789-797.

Sheen P, Couvin D, Grandjean L, Zimic M, Dominguez M, Luna G, Gilman RH, Rastogi N, Moore DA. 2013. Genetic diversity of *Mycobacterium tuberculosis* in Peru and exploration of phylogenetic associations with drug resistance. *PLoS One*. 8(6): e65873.

Smit PW, Haanperä M, Rantala P, Couvin D, Lyytikäinen O, Rastogi N, Ruutu P, Soini H. (2013). Molecular epidemiology of tuberculosis in Finland, 2008-2011. *PLoS One*. 8(12): e85027.

Smith NH, Upton P. 2012. Naming spoligotype patterns for the RD9-deleted lineage of the *Mycobacterium tuberculosis* complex; www.Mbovis.org. *Infect Genet Evol*. 12 (4), pp. 873–876.

Soares P, Alves RJ, Abecasis AB, Penha-Gonçalves C, Gomes MG, Pereira-Leal JB. (2013) inTB – a data integration platform for molecular and clinical epidemiological analysis of tuberculosis. *BMC Bioinformatics*. 14: 264.

Sola C, Devallois A, Horgen L, Maïsetti J, Filliol I, Legrand E, Rastogi N. 1999. Tuberculosis in the Caribbean: using spacer oligonucleotide typing to understand strain origin and transmission. *Emerg Infect Dis*. 5(3): 404-14.

Sola C, Filliol I, Gutierrez MC, Mokrousov I, Vincent V, Rastogi N. 2001. Spoligotype database of *Mycobacterium tuberculosis*: biogeographic distribution of shared types and epidemiologic and phylogenetic perspectives. *Emerg Infect Dis*. 7(3): 390-396.

Sola C, Filliol I, Legrand E, Mokrousov I, Rastogi N. 2001. *Mycobacterium tuberculosis* phylogeny reconstruction based on combined numerical analysis with IS1081, IS6110, VNTR, and DR-based spoligotyping suggests the existence of two new phylogeographical clades. *J Mol Evol*. 53: 680-689.

Sola C, Filliol I, Legrand E, Lesjean S, Locht C, Supply P, Rastogi N. 2003. Genotyping of the *Mycobacterium tuberculosis* complex using MIRUs: association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics. *Infect Genet Evol*. 3: 125-133.

van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, Portaels F, Qing HZ, Enkhsaikan D, Nymadawa P, van Embden JD. (1995). Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. *J Clin Microbiol.* 33: 3234-3238.

Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS, Musser JM. 1997. Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci USA*. 94: 9869-9874.

Supply P, Lesjean S, Savine E, Kremer K, van Soolingen D, Locht C. 2001. Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. *J Clin Microbiol*. 39: 3563-3571.

Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E, Savine E, de Haas P, van Deutekom H, Roring S, Bifani P, Kurepina N, Kreiswirth B, Sola C, Rastogi N, Vatin V, Gutierrez MC, Fauville M, Niemann S, Skuce R, Kremer K, Locht C, van Soolingen D. 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. *J Clin Microbiol*. 44: 4498-4510.

Tang C, Reyes JF, Luciani F, Francis AR, Tanaka MM. 2008. SpolTools: online utilities for analyzing spoligotypes of the *Mycobacterium tuberculosis* complex. *Bioinformatics*. 24: 2414-415.

Wolfe ND, Dunavan CP, Diamond J. 2007. Origins of major human infectious diseases. *Nature*. 447: 279-283.

Zink AR, Sola C, Reischl U, Grabner W, Rastogi N, Wolf H, Nerlich AG. 2003. Characterization of *Mycobacterium tuberculosis* complex DNAs from Egyptian mummies by spoligotyping. *J Clin Microbiol*. 41: 359-367.

Weniger T, Krawczyk J, Supply P, Niemann S, Harmsen D. 2010. MIRU-VNTRplus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. *Nucleic Acids Res*. 38(Web Server issue): W326–W331.

69.WHO. 2006. The Stop TB Strategy: Building on and enhancing DOTS to meet the TB-related Millennium Development Goals. http://whqlibdoc.who.int/hq/2006/WHO_HTM_STB_2006.368_eng.pdf?ua=1.

WHO. 2013. Global tuberculosis report 2013.

www.who.int/iris/bitstream/10665/91355/1/9789241564656_eng.pdf

EuroReference
journal of Reference

anses
French agency for food, environmental and occupational health & safety

Journal No. 12
Summer 2014

Agenda

Publication and design
Publication director: Marc Mortureux
Editor-in-chief: Paul Martin

Creation/Development: Julien Vigneron, Céline Leterq, Fabrice Coutureau, Parimage
ISSN 2110-5294

French Agency for Food, Environmental
and Occupational Health & Safety
27-31 avenue du général Leclerc
F-94701 Maisons-Alfort Cedex - France
www.anses.fr