Development of a miniaturized DNA microarray for identification of 66 virulence genes of *Legionella pneumophila*

Opracowanie miniaturowej mikromacierzy DNA do identyfikacji 66 genów wirulencji *Legionella pneumophila*

Mariusz Żak¹, Piotr Zaborowski², Milena Baczewska-Rej³, Aleksandra A. Zasada³, Renata Matuszewskaa³, Bożena Krogulska⁴

¹ Department of Zoonoses and Tropical Diseases, Medical University of Warsaw, Żwirki i Wigury 61 Str., 02-091 Warsaw, Poland
² Department of Internal Diseases, Gastroenterology and Hepatology, University Hospital, University of Warmia and Mazury in Olsztyn, Oczapowskiego 2 Str., 10-719 Olsztyn, Poland
³ Department of Bacteriology, National Institute of Public Health – National Institute of Hygiene. Chocimska 24 Str., 00-791 Warsaw, Poland
⁴ Department of Communal Hygiene. National Institute of Public Health – National Institute of Hygiene, Chocimska 24 Str., 00-791 Warsaw, Poland

**Summary**

**Introduction:**
For the last five years, *Legionella* sp. infections and legionnaire’s disease in Poland have been receiving a lot of attention, because of the new regulations concerning microbiological quality of drinking water. This was the inspiration to search for and develop a new assay to identify many virulence genes of *Legionella pneumophila* to better understand their distribution in environmental and clinical strains. The method might be an invaluable help in infection risk assessment and in epidemiological investigations.

**Material/Methods:**
The microarray is based on Array Tube technology. It contains 3 positive and 1 negative control. Target genes encode structural elements of T4SS, effector proteins and factors not related to T4SS. Probes were designed using OligoWiz software and data analyzed using IconoClust software. To isolate environmental and clinical strains, BAL samples and samples of hot water from different and independent hot water distribution systems of public utility buildings were collected.

**Results:**
We have developed a miniaturized DNA microarray for identification of 66 virulence genes of *L. pneumophila*. The assay is specific to *L. pneumophila* sg 1 with sensitivity sufficient to perform the assay using DNA isolated from a single *L. pneumophila* colony. Seven environmental strains were analyzed. Two exhibited a hybridization pattern distinct from the reference strain.

**Discussion:**
The method is time- and cost-effective. Initial studies have shown that genes encoding effector proteins may vary among environmental strains. Further studies might help to identify set of genes increasing the risk of clinical disease and to determine the pathogenic potential of environmental strains.

**Key words:** DNA microarray • virulence factor • effector protein • *Legionella pneumophila* • legionellosis
Streszczenie

Wprowadzenie: Przez ostatnie pięć lat coraz więcej uwagi w Polsce zaczęto poświęcać problemowi infekcji wywoływanych przez bakterie Legionella sp. Ma to związek m.in. z pojawieniem się nowych regulacji prawnych, dotyczących jakości wody przeznaczonej do spożycia przez ludzi. Było to inspirażją do opracowania testu umożliwiającego oznaczenie wielu genów wirulencji Legionella pneumophila w celu lepszego zrozumienia ich dystrybucji w szczepach środowiskowych i klinicznych. Metoda może być nieocenioną pomocą w skutecznym przewidywaniu ryzyka wywołania choroby przez dany szczep i stanowić nowe narzędzie podczas dochodzeń epidemiologicznych.

Materiał/Metody: Mikromacierz bazuje na systemie Array Tubes, zawiera trzy kontrole pozytywne i jedną negatywną. Geny docelowe kodują elementy strukturalne T4SS, białka efektorowe oraz geny wirulencji niezwiązane z systemem sekrecji. Sondy zaprojektowano używając programu OlogoWiz, mikromacierz analizowano z użyciem programu IconoClust. W celu izolacji szczepów klinicznych i środowiskowych zgromadzono próbki BAL i pobrano próbki wody z różnych, niezależnych instalacji ciepłej wody w budynkach użyteczności publicznej.

Wyniki: Zaprojektowano i wyprodukowano mikromacierz do identyfikacji 66 genów L. pneumophila biorących udział w patogenezie. Czułość mikromacierzy pozwala na analizę DNA wyizolowanego z pojedynczej kolonii L. pneumophila. Analizowano 7 szczepów środowiskowych, z których dwa różniły się wzorem hybrydyzacji od szczepu referencyjnego.

Dyskusja: Prezentowana metoda nie jest czasochłonna i kosztowna. Analiza szczepów środowiskowych wykazała możliwe różnice w obecności genów kodujących białka efektorowe. Dalsze analizy mogą pozwolić na identyfikację genów zwiększających ryzyko wywołania choroby oraz na ocenę zjawistości szczepów.

Słowa kluczowe: mikromacierz DNA • czynnik wirulencji • białko efektorowe • Legionella pneumophila • legioneloza

Acknowledgements: mgr Mariusz Żak, ul. Niepodległości 17/22, 05-600 Grójec; e-mail: mariuszakster@gmail.com

Abbreviations: AFLP – amplified fragment length polymorphism; BAL – broncho-alveolar lavage fluid; BLAST – basic local alignment search tool; EWGLI – European working group for Legionella infections; LCV – Legionella containing vacuole; MLVA – multiple-locus variable number of tandem repeats analysis; NSI – normalized signal intensity; PCR – polymerase chain reaction; PFGE – pulsed-field gel electrophoresis; SBT – sequence-based typing; T4SS – type IV secretion system.
After infection the disease may develop up to 20 days. This means that identification of the source of infection is difficult, especially if the disease developed after travel. To establish the link between environmental and clinical strains, genetic studies are required. Routinely PFGE, considered as a gold standard, SBT, AFLP and MLVA are used. These techniques take advantage of identification of DNA polymorphism after restriction enzymes digestion (PFGE, AFLP), analysis of multiple locus variable number tandem repeats (MLVA) and sequencing of selected genes (SBT) [21,22,41,51,60]. The methods have different ability to discriminate strains based on the presence in the genome of DNA polymorphisms unspecific for virulence, but do not allow particular virulence genes to be identified. The technique which allow this are microarrays. The method has been used for the rapid identification of antimicrobial resistance genes in Gram-negative bacteria [6], genotyping of enterohemorrhagic Escherichia coli (EHEC) [24] and Staphylococcus aureus [32], detection of herpesvirus and adenovirus co-infections [38] and for haemagglutinin subtyping and pathotyping of avian influenza viruses [23]. Simultaneous identification of many virulence genes is essential, because effector proteins of L. pneumophila display a high level of redundancy. Deletion of a single gene often does not cause a noticeable effect on phenotype [17]. Identification of many virulence genes in environmental strains will give a more reliable view on the virulence potential of a given strain and will help to better understand their distribution. In future it may help to predict with higher accuracy risk of infection.

Risk assessment of infection is usually made on the basis of the number of Legionella sp. in tested water. Routinely, for quantitative analysis, membrane filtration and cultivation on solid microbiological media is performed. The method in detail is described in PN – EN ISO 11731-2: 2008 [45]. In Polish law a regulation that relates directly to this issue is the decree of Ordinance of the Minister of Health of March 29, 2007 on the requirements related to the quality of water intended for human consumption [47,48]. European guidelines, formed by EWGLI in 2005, contain many instructions about risk assessment and management in order to minimize risk of infection [27]. Attachment No. 7 to the Polish decree defines the risk based on the number of Legionella sp. determined according to PN – EN ISO 11731-2: 2008. The risk gradation is consistent with the EWGLI guidelines, but the regulation says nothing about risk assessment and management performed individually for each installation. The assessment of risk of infection based on the number of Legionella sp. is very difficult, because infection dose is still not determined and not all individuals exposed will develop the disease.

To date, five genomes of L. pneumophila have been sequenced. Pan-genome analysis reveals that only 67% of identified genes are common for all strains, while 33% are strain-specific [15]. Such high genetic diversity might be very useful for pathotyping of a strain.

This was the inspiration for our study, in which we present a low-density oligonucleotide DNA microarray dedicated to identification of 66 virulence genes of Legionella pneumophila sg 1. Genes of interest are involved in translocation of effector proteins into the eukaryotic cell ( dot/icm genes encode structural elements of T4SS, and proteins with function similar to molecular chaperones), growth on solid microbiological media, intracellular growth, inhibition of apoptosis, maturation of LCV, iron acquisition, and cytochrome c synthesis. Function of many target genes encoding effector proteins is not determined. Target genes encoding effector proteins were selected on the basis of studies where their translocation to a eukaryotic cell was proven. Target genes are listed in Table 1. All probes and primers were designed in this study.

**Materials and Methods**

**Microarray design and manufacture**

The designed microarray is based on the Array Tube system (Alere Technologies GmbH). On the surface 9 mm2 of glass slide probes are immobilized via amino-linker C7. Each slide has its own hybridization chamber. In the first step of the detection procedure, products of linear PCR, labeled with biotin, hybridize to specific probes. During the next step, streptavidin conjugated with horseradish peroxidase binds to the biotin. The enzyme transforms the soluble and colorless substrate into an insoluble black product. In the last step each microarray is analyzed in an array reader and clarity of all spots is calculated.

All probes were designed using OligoWiz 2.0 software. Legionella pneumophila subsp. pneumophila str. Philadelphia 1 genome was used as a reference sequence. Probes were designed in two stages. In the first stage, a set of probes specific to each target gene was designed (up to 15 probes for each gene). Probe sequences were selected to be specific for the target gene of the reference strain and to have specified GC contents, lengths, and Tm in order to yield high hybridization efficiency. In the second stage the BLAST algorithm was used to select one probe specific to the consensus region of the target gene.

Primers were designed using Primer-BLAST. Primers were selected to be specific and to have similar GC contents, lengths, and Tm in order to have similar template binding efficiency. Probes and primers are specific to a consensus region of target genes. The primer binding site is situated up to 100 bp upstream of the probe-binding site (Figure 1). Probes and primers are listed in Table 2.

Each target gene has one specific probe and each spot is duplicated. Spot distance: 180 µm, spot diameter: 150 µm (Figure 2). The microarray has one negative control (NC) – a probe and primer complementary to human papilloma virus E7 gene – and 3 positive controls – probes and primers complementary to the gene encoding 16S rRNA (lpg0302), mip gene (lpg0791) encoding macrophage infectivity potentiator, and pgI gene (lpg0759) encoding glucose-6-phosphate isomerase.

**Bacterial strains**

Legionella pneumophila was isolated from water samples collected from the hot water systems according to PN – EN ISO 11731-2: 2008. Serogroup 1 was identified using an L. pneumophila latex kit (BIORAD) according to the manufacturer’s instructions. (Five strains were obtained from the Department of Communal Hygiene, National Institute of Public Health – National Institute of Hygiene, Warsaw. Two
Table 1. Genes of interest

| No. | Effector proteins                  | 36 | VpdA | lpg2410 | Reference |
|----|-----------------------------------|----|------|--------|-----------|
| 1  | LegA11                            |    |      |        |           |
| 2  | LegA12                            |    |      |        |           |
| 3  | LegA15                            |    |      |        |           |

| No. | Structural elements of T4SS          | 37 | VpdB | lpg1227 | Reference |
|----|-------------------------------------|----|------|--------|-----------|
| 3  | LegA15                              |    |      |        |           |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 1  | LegA11  | lpg0436| [18]      |
| 2  | LegA12  | lpg0436| [18]      |
| 3  | LegA15  | lpg2456| [18]      |

| No. | Other virulence factors          | 38 | WipA | lpg2718 | Reference |
|----|----------------------------------|----|------|--------|-----------|
| 1  | LegA21                            |    |      |        |           |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 38 | WipB    | lpg0642| [33]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 39 | WipB    | lpg0642| [33]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 40 | DotB    | lpg2676| [57]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 41 | DotC    | lpg2675| [54]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 42 | DotD    | lpg2674| [54]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 43 | icmB(DotO) | lpg0456| [1]       |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 44 | icmC(DotE) | lpg0453| [54]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 45 | icmF    | lpg0458| [52]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 46 | icmG(DotF) | lpg0452| [54]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 47 | icmH(DotI) | lpg0459| [52]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 48 | icmI(DotN) | lpg0455| [54]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 49 | icmJ(DotJ) | lpg0448| [1]       |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 50 | icmK(DotH) | lpg0450| [54]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 51 | icmL(DotL) | lpg0449| [1]       |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 52 | icmM(DotM) | lpg0445| [54]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 53 | icmN(DotK) | lpg0446| [31]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 54 | icmO(DotL) | lpg0447| [58]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 55 | icmP(DotM) | lpg0448| [50]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 56 | icmQ    | lpg0444| [50]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 57 | icmR    | lpg0443| [50]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 58 | icmT    | lpg0441| [35]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 59 | icmW    | lpg2688| [9]       |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 60 | icmX    | lpg1284| [36]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 61 | CcmC    | lpg0858| [56]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 62 | IraA    | lpg0747| [55]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 63 | IraB    | lpg0746| [55]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 64 | LetA    | lpg0747| [16]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 65 | Mip     | lpg1284| [36]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 66 | RpoS    | lpg0103| [49]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 67 | VpdD    | lpg2831| [49]      |

| No. | Protein | Gen ID | Reference |
|----|---------|--------|-----------|
| 68 | VpdF    | lpg0103| [49]      |
**Fig. 1.** Principle of the method. Location of hybridization sites of probes and primers on a DNA template. Arrow indicates the direction of elongation.

**Table 2.** Names and sequences of all designed probes immobilized on the microarray and primers used in linear multiplex PCR

| Probe name | Sequence 5’-3’ | Primer name | Sequence 5’-3’ |
|------------|----------------|-------------|----------------|
| ccmC       | TTGATGCTCGGAGGCTGGTCAGTTGG| ccmC-p       | AGGAGCCCGATATTTTGGAGTACG |
| dotB       | TTAGCGGAGAGAGCGTCGTTGGGAGAAGA | dotB-p | GGCAAGGCGAAGGACTTATTGATG |
| dotC       | GCGAGATGCTGGCTTAAAGGACCAGGGCCG | dotC-p | GCTCGAGTTTTAATGTGGTCGTC |
| dotD       | TGGATGAGACGAGCGTCTGGTGTGGTT | dotD-p | CGTAAATGCTGGTGTGGTCG |
| icmB(dotO) | GGCAATCCTGGGAAAGTGACGTACG | icmB(dotO)-p | GTCACCCATGGGAACAGTGAC |
| icmC(dotE) | GCTGATCAGCTGATGCTGGGACGAC | icmC(dotE)-p | GTTGAAGGCTCGCTACAAATG |
| icmD(dotP) | AGACGAGCTGGCTCTACTGCTGATATGGTT | icmD(dotP)-p | GTGGATATCGCTGCTGGTTAAA |
| icmE(dotG) | GGAAGGCGATGGATCACAGGGCCG | icmE(dotG)-p | CAATCATTACACCTGTAACCAGA |
| icmF       | TTAGATGATGACTAATGCCAGGGCTCTG | icmF-p | CCAATTTAGATCCGTGCAGACTGAGA |
| icmG(dotD) | AATTCAGGCTGGCTAGTTG | icmG(dotD)-p | GGGGATATTTTGGACCTTC |
| icmH(dotU) | ACCAATGAGAGGAGGCTGGCAGACTG | icmH(dotU)-p | GGGATGCTGGTATGAAACG |
| icmI(dotN) | TGAATGAGACTAGCTAATGAGGCTTCTG | icmI(dotN)-p | AATACCATGACGAGCTCCAG |
| icmK(dotH) | ATATTTGAGCTGGGATGTTGAGGCTG | icmK(dotH)-p | TAACTGATGCTGTGCTGTC |
| icmL(dotO) | GGCTGATTTGAGCTGGGACGAGGCTG | icmL(dotO)-p | TTGCTGCTTTAACGAGCTCAGA |
| icmM(dotU) | ATTCGGACGAGGGCAATCCTTCCCTC | icmM(dotU)-p | CAATGCTTGAGATGCTAGTGAAG |
| icmO(dotL) | GTATGCTGGACGAGGCTGGTCTTCTC | icmO(dotL)-p | TGTCAAGGGAAATACAGCAGA |
| icmP(dotM) | GCTGATGGCGGCTTGATTG | icmP(dotM)-p | AAGAAGAAGCGGACAGTACG |
| icmQ       | AAGGGATAGTGCTGGCGCTGACG | icmQ-p | TGCAATCTGGAAGATGACTCTG |
| icmR       | TGAATGAGACTAGCTAATGAGGCTTCTG | icmR-p | CCTATCTCTTACAGAGGAGCTG |
| icmT       | CCTGCGGCTGAGTCTGGGAGAGGAGAAG | icmT-p | AGAGAAAAACAAATGGGAGGAGG |
| icmW       | GCTGGATTTAGCCGTAGCAAGGCGCCC | icmW-p | TGATGGACCTGAAATACCTG |
| pgi        | TGAATGAGACTAGCTAATGAGGCTTCTG | pgi-p | AGGGATACAGCAGTACG |
| legA2      | GGCCTTATCGAGGCTGGCTGGAGCAG | legA2-p | TCCTTCTGGAAGAAAAAGAGAAG |
| legA3      | TGATCTTCTCCTGGGCGAGTCTG | legA3-p | ACTGGTGAACGAGGAGAAG |
| legA8      | TGGAGATGCTAATGAGGCTTTCTTCTC | legA8-p | GGGATGAAATCTTCTGAGTAA |
| legA      | TGGAGATGCTAATGAGGCTTTCTTCTC | legA-p | GGGATGAAATCTTCTGAGTAA |
| legB      | TAAAGGCGAGGCGCTGCCGACG | legB-p | TGATGCGATCAGGAGTATTAG |
| lidA       | ATGCAGAAATCTGGCTGGAGCTG | lidA-p | AAGTGAATGCAGGAGTATTAG |
| dotK       | CCGTGGATGCTAGCTGGGAGAG | dotK-p | TGGTGATAGCTGGGAGTATTAG |
| mip        | GCGATGAGACTAGCTGAGGCTG | mip-p | AAGCGAGTACCGAGTACG |
| rpoS      | GCGCTGACGAGGCGCTGGAGGAGCAG | rpoS-p | CCACCCGGCAGCTACAAAGAGT |
| sdhA       | TGAAGAACGAGGCTGGGAGGCTG | sdhA-p | AAGTGGTACAGTCTGGAGT |
| sidG       | AAGGCGAGGCGCTGGGAGGCTG | sidG-p | GGGATGAAATCTTCTGAGTAA |
| sidH       | ATGGCATGAGGCTGGAGGCTG | sidH-p | TGGTGATAGCTGGGAGTATTAG |
| sidJ       | GCACTGATGACGCTGGAGGCTG | sidJ-p | AGGCCAACCTAATACCT |
| vipA       | AATATGCGGCGAATGCGGAGG | vipA-p | TGATGCGATCAGGAGTATTAG |
strains were isolated in this study.) *Legionella pneumophila* sg 1 Philadelphia 1 (ATCC 31152) was used as a reference strain. In order to isolate clinical strains of *L. pneumophila*, BAL samples were collected from 31 patients with chronic cough, hemoptysis, chest pain, diarrhea, respiratory disorders, abdominal pain, confusion. Samples were cultivated on BCYE-a with cysteine and confirmed on BCYE-a without cysteine and BMPA plates. Not a single strain was isolated. To confirm the negative result of the cultivation, DNA from BAL samples was isolated and used as a template for real time PCR reaction to detect and quantify *Legionella* sp. (BioRad). In all tested samples *Legionella* sp. DNA was not detected.

Table 2 continued. Names and sequences of all designed probes immobilized on the microarray and primers used in linear multiplex PCR

| Probe name | Sequence 5’-3’ | Primer name | Sequence 5’-3’ |
|------------|----------------|-------------|----------------|
| vipD       | CCTGTGCAAAGGCTCGGCAAGA | vipD-p | TCTACTTTGTCAGAGTTGGAAAC |
| vipF       | TTGAGGGAAGGCAATTCGCGGT | vipF-p | TTTTGAAATAGAATCAGCCGAA |
| irA        | AAATCTCTGATCTGGGCTGTC | irA-p | CATCACTATGGCAATTTGGCA |
| letA       | TGGCTGGGAGTATGCGGGAGAAGTG | letA-p | GAGAGGATGTCGGCAAAAATAGCT |
| legC8      | TGGCAAAGCAGCTGCTGGTCGT | legC8-p | GATTGTTGACGCTCAACTTC |
| vpdA       | CAAATACCTGAGGCAATGCCCATA | vpdA | GATTGGTCAAAAGCAATAGTTGG |
| vpdB       | AGGAGTGGACAGCTGAGGATGG | vpdB-p | CAGAGATTGCAAAGAATGGTG |
| wipA       | TCACTGGGGACTTGTGATGCGGCAAT | wipA-p | TGGAGGAAAGAAATAAGGTAAAG |
| wipB       | CAGGACAGCTACTGGGGACCAAGG | wipB-p | GATTCCTTTAAACGCCTGGT |
| legA11     | GCCCCCTATTACGACAAAGCAGCAAC | legA11-p | TGATTATGCTGCTTACACCAAC |
| legA12     | TATATTCGCCCCTGCGATCAACC | legA12-p | GACCTCTTTTAAGAAAGCAACC |
| legA54     | AATCTCTGACGCTGTCGTCAGCGG | legA54-p | AGGCTATTCAGGCAACCCAAG |
| legAU13    | TGGGAAAATCCATTGATGCGGAGGCA | legAU13-p | TTGACCTCTCCAGGTTATTC |
| legC3      | GGGTTAAAGGCGCGCAAGAAATGG | legC3-p | GAATAGCTGGCAAAATAGGG |
| legC4      | TGGGAAATACCTGAGGCAAGCAG | legC4-p | CATGTGGGCGATATTGTTT |
| legC5      | GGGTTTCCAGAATCTACCAAGGATGTCT | legC5-p | TGGAAATTACTACGCGCCAAA |
| legC6      | TGGCTCTCATGATGCGGACAGCGG | legC6-p | GAACGTGCGAAGGATTTT |
| legC7      | TGGGACAGCCACACATACGCTAAGTGA | legC7-p | CCCCCTTCTTTATGTTGGGCA |
| legG1      | ATGCCCATAATGGGATGGCGGAGGATA | legG1-p | AGTGGGACAGATCTACTAG |
| legG2      | CGGGATGACAGCGGACCAGCTGCTTGG | legG2-p | GCTCAGCCATCAATTTTAC |
| legL2      | GGCAGGTCTGCTGATGTCACGCAAAGG | legL2-p | ATCAGTTGCAATGGGTCTCC |
| legL3      | AATGCAAAAAGAATGGGCCATGCTG | legL3-p | GATAGAGGCGCGGCTGATAATC |
| legL6      | GGGAGCCATATCCTCTGCAGCCAC | legL6-p | TCTGGACGCTGCTTTATGCTG |
| legL7      | AGGACGGCATGCTGATGCGGACAGCGG | legL7-p | AACTCCAGAATTGTTAGG |
| legL8      | TGGTATGAGGCGGATCCGCAAGGCTGTC | legL8-p | CCCCCTTACTTACCATACCA |
| legL9      | CATTGCAGCGCAGGCAAGAGCT | legL9-p | CTTCTACTTTATGCTGCTTC |
| legL10     | GAAATAGCAGCGGACCTGCTGGGCTTGC | legL10-p | GAGCTGCGTATTGTTGCTCAAG |
| legL11     | CGGGTTCGCGCAAGGGCAAAAAGAATA | legL11-p | TGGGAGAATGACAGGTTT |
| legL12     | ATATTGCTCGGGAGGCTGCTAGCCATG | legL12-p | CTTGTCACCCAAATCCTGGT |
| legL13     | TGTTGACGCGGCAAAAAGGCT | legL13-p | GAGGCTATTGTCGAGGTT |
| legL14     | CGGAGGCGGCTGCTGCTTCGCTGTC | legL14-p | GCAATAAGGATAAGGGTGGG |
| NC         | TTTAAGGGGAACGTCCCGGCAAGT | NC-p | GAAGAGGCTGGGCAAGGAA |

BCYE-α with cysteine and confirmed on BCYE-α without cysteine and BMPA plates. Not a single strain was isolated. To confirm the negative result of the cultivation, DNA from BAL samples was isolated and used as a template for real time PCR reaction to detect and quantify *Legionella* sp. (BioRad). In all tested samples *Legionella* sp. DNA was not detected.
DNA isolation and preparation

One colony of each strain of *L. pneumophila* was resuspended in lysis buffer from DNeasy Blood & Tissue Kit and the DNA was isolated according to the manufacturer’s instructions. BAL samples (5 ml) were centrifuged (10 min., 12000 rpm), the pellet was resuspended in lysis buffer from DNeasy Blood & Tissue Kit and the DNA was isolated according to the manufacturer’s instructions. DNA concentration was measured using a spectrophotometer. DNA isolated from bacterial strains was subsequently diluted in double-distilled water to the expected concentration. DNA isolated from BAL samples was concentrated to the expected concentration using BlueMATRIX PCR/DNA Clean-up Purification Kit (Eurex) according to the manufacturer’s instructions.

Labeling and PCR

Each elongation reaction contained 1 µl dNTP mix (Fermentas), 1 µl Terminator DNA polymerase buffer 10× (New England Biolabs), 0.1 µl Terminator DNA polymerase (New England Biolabs), 0.35 µl Biotin-16-dUTP (Roche), 1 µl primer mix (Genomed), and 1 µg genomic DNA (up to 6.55 µl). The PCR conditions were as follows: 3 min at 96°C, followed by 40 cycles of 20 s at 51°C, 40 s at 72°C and 60 s at 96°C.

Hybridization and detection

Each Array Tube was in the first step washed with 500 µl of double-distilled water and incubated at 55°C using a thermodemixing device for 5 min at 550 rpm. In the next step the washing buffer was discarded and 500 µl of hybridization buffer was added for 5 min at 55°C and 550 rpm. The hybridization buffer was removed from the hybridization chamber and 100 µl of a hybridization mix (90 µl of hybridization buffer and 10 µl of PCR product) was added for 60 min at 55°C and 550 rpm. In the next step the Array Tube was washed once with 500 µl of 2 × SSC with Triton X-100 (10 × SSC is 1.5 M NaCl, 0.15 M sodium citrate, pH 7.0) for 5 min at 30°C and 550 rpm, once with 500 µl of 2 × SSC for 5 min at 30°C and 550 rpm and finally with 500 µl of 0.2 SSC for 5 min at 20°C and 550 rpm. The washing buffer was discarded and 100 µl of blocking solution was added (blocking solution is 0.005% Triton X-100 and 60 mM sodium phosphate, 1.08 M NaCl, 6 ml EDTA, pH 7.4) for 15 min at 30°C and 550 rpm. Blocking solution was removed and horseradish peroxidase-streptavidin conjugate solution was added (Poly-HRP-streptavidin was diluted 1: 3000 in 6 × SSPE (6 × SSPE contains 0.005% Triton X-100 and 0.01% sodium phosphate, 1.08 M NaCl, 6 ml EDTA, pH 7.4)) for 15 min at 30°C and 550 rpm. To remove excess HRP, the Array Tubes were washed using washing buffers as described above. In the last step, 100 µl of peroxidase substrate was added (Seranum Grün, Seranumagnostica GmbH) and the staining reaction was performed at ambient temperature without shaking for 10 min. Array Tubes were placed in an Array Tube reader (Alere Technologies GmbH). Data were acquired and analyzed using IconoClust 3.2 software (Alere Technologies GmbH).

Data analysis

Analysis was performed using IconoClust 3.2 software. Raw data were normalized and the Normalized Spot Intensity (NSI) was calculated. NSI = 1 - (M/BG) where M represents the average intensity of a spot and the BG represents the intensity of the local background. NSI value >0.1 of a given spot was considered to be “positive”.

Results

Assay validation

In the first step of the validation, quality control of the designed and manufactured microarrays was performed. To exclude contamination of probes with biotin, the full
hybridization procedure was performed using double-distilled water in place of biotinylated PCR product. NSI of spots corresponding to the legL1 gene was higher in comparison with background and other spots on the microarray. In further analysis these spots were not considered because of possible contamination with biotin. All data were normalized to minimize the effect of heterogeneous background signal emission (Figure 3D).

Specificity of the microarray was tested using genomic DNA isolated from *Legionella pneumophila* subsp. *pneumophila* str. Philadelphia 1, *Escherichia coli* (O157: H7), *Pseudomonas aeruginosa*, *Haemophilus influenzae* and DNA isolated from BAL samples. The PCR product used for hybridization, where DNA of *Escherichia coli* (O157: H7), *Pseudomonas aeruginosa* and *Haemophilus influenzae* was the template, hybridized only to the probes corresponding to 16S rRNA (Figure 3A). The PCR product from BAL samples gave no positive signal in any spot area (Figure 3C). The PCR product from the *L. pneumophila* reference strain hybridized to all probes corresponding to target genes. The spot corresponding to the icmL probe characterized low value of the NSI in all tested microarrays. Hybridization to probes corresponding to negative control was not detected (Figure 3B).

**Assay sensitivity**

To determine the minimal DNA concentration for the analysis, a series of dilutions of reference DNA was made: 250 ng/µl, 150 ng/µl, 100 ng/µl and 50 ng/µl. The detection limit was 975 ng DNA/reaction. It is the equivalent of 2.66x10^8 DNA copies, based on a genome size of 3.4x10^6 nucleotide residues corresponding to a 3.6x10^-15 g/1 genome copy. In experiments with DNA dilutions near to the detection limit, heterogeneous signals from different spots were observed. To overcome this effect, 1.2 µg of DNA should be used in routine experiments.

**Tested strains**

In the last stage of the study we tested DNA isolated from 7 environmental strains. Analysis of DNA isolated from five strains showed a positive signal in all spots corresponding to target genes. DNA from two tested strains, named LPE06 and LPE07, gave a spot signal pattern different to the reference strain and the other environmental strains (Figure 4). Exact analysis of spot signal level indicated that both strains give a positive signal of spots corresponding to dot/icm genes. Strain LPE06 gives no signal of spots corresponding to legC8, legLC4, legU2, sidG, SidH and wipB genes. Strain LPE07 gives no signal of spots corresponding to legC12, legC4, legC8, legL7, legLC4, legU1, legU2, lepB, sidG, sidH, vipD and wipB genes.

**Discussion**

Routinely, legionellosis risk assessment is made on the basis of the number of *Legionella* sp. detected in a water sample using microbiological methods. Results of the analysis determine actions which should be put into practice to reduce the level of contamination. These actions are often expensive and cause difficulties in usage of the installation. If dead ends are present, effective chemical or thermal disinfection is very difficult to perform. The method also has limitations. Identification of *L. pneumophila* in water from cooling towers has shown that the number of bacteria may vary in a short period of time. Single water testing may not give reliable information about the level of contamination of installation [7]. Other studies have
shown that genotypes of *L. pneumophila* in cooling towers do not change while the number of bacteria vary [46]. For this reason, the bacteria number count should not be the only method for the assessment of risk of infection. The microbiological method should be supplemented by analysis of virulence potential of a strain, which is determined by specific virulence genes.

The presented microarray allowed identification of two strains with a hybridization pattern different to the reference strain. Strain LPE07 does not give a positive signal of spots corresponding to genes encoding 6 effector proteins and strain LPE07 to 12 effector proteins. Previous studies with microarrays have shown that genes *sidG* and *sidH* are not highly conserved. Gene *sidG* was present in 52% of tested strains while *sidH* was present in 72%. Genes encoding structural elements of T4SS were present in all tested strains except the gene *icmX*, which was present in 65% of tested strains [11]. These data are compatible with the results of our study, because all spots corresponding to dot/icm genes give a positive signal. Strains LPE06 and LPE07 give no positive signal for three and seven eukaryotic-like genes respectively. Leg genes have been discovered recently and their function in many cases is still not known. Further analysis of clinical and environmental strains isolated from different installations made of different materials with parallel identification and environmental strains isolated from different installations made of different materials with parallel identification of protozoa will help to better understand their distribution and functions in pathogenesis. Highly conserved virulence genes are important for survival of *Legionella*. Strains with deletions in these genes may have reduced ability to infect protozoa and therefore to survive in water with biofilms and to infect humans. Knowledge about the virulence potential of a strain identified in an installation of interest with information about the number of bacteria will help to choose the most effective intervention.

Sensitivity of the array allows one to analyze DNA isolated not only from one colony and is similar to sensitivity levels reached in other studies with this technology [2,4,37]. Furthermore, there are several advantages of this technology. In one experiment detection of up to 100 genes in duplicate is possible in only 8 hours. It is less costly and time-consuming than 200 PCR reactions. The Array Tube system allows one to perform the assay with standard laboratory equipment, because each Array Tube has the Eppendorf tube format. Moreover, PCR is prone to contamination, which is almost impossible with the Array Tube system, and a free PCR product environment is not rigorously required. It is an open system and changes in microarray layout are easy to introduce. In future the method might be a supplementation of the microbiological method in routine diagnostics of environmental samples.

Studies performed at the National Institute of Public Health – National Institute of Hygiene showed that the bacteria were present in most tested samples collected from hot water distribution systems in inpatient healthcare facilities [34]. Moreover, no method of disinfection applied once guarantees permanent elimination of the pathogen. Only constant monitoring will give valid information about the risk of infection. The microbiological method has limitations and because of that the microarray was designed. The technique has been used in many fields of environmental and medical diagnostics and it is also a promising method for identification and graduation of risk of infection of *Legionella pneumophila*.

**References**

[1] Andrews H.L., Vogel J.P., Isberg R.R.: Identification of linked *Legionella pneumophila* genes essential for intracellular growth and evasion of the endocytic pathway. Infect. Immun., 1998; 66: 950–958

[2] Anjum M.F., Mafura M., Slickers P., Ballmer K., Kuhnert P., Woodward M.J., Ehrlich R.: Pathotyping *Escherichia coli* by using miniaturized DNA microarrays. Appl. Environ. Microbiol., 2007; 73: 5692–5697

[3] Bachman M.A., Swanson M.S.: Genetic evidence that *Legionella pneumophila* RpoS modulates expression of the transmission phenotype in both the exponential phase and the stationary phase. Infect. Immun., 2004; 72: 2468–2476

[4] Ballmer K., Korczak B.M., Kuhnert P., Slickers P., Ehrlich R., Hachler H.: Fast DNA serotyping of *Escherichia coli* by use of an oligonucleotide microarray. J. Clin. Microbiol., 2007; 45: 370-379

[5] Bandill J.P., Miller J.L., Vogel J.P.: IcmS-dependent translocation of SdeA into macrophages by the *Legionella pneumophila* type IV secretion system. Mol. Microbiol., 2005; 56: 90–103

[6] Batchelor M., Hopkins K.L., Liebana E., Slickers P., Ehrlich R., Mafura M., Aarestrup F., Mevius D., Clifton-Hadley F.A., Woodward M.J., Davies R.H., Thriffall E.J., Anjum M.F.: Development of a miniaturized microarray-based assay for the rapid identification of antimicrobial resistance genes in Gram-negative bacteria. Int. J. Antimicrob. Agents, 2008; 31: 440–451

[7] Bentham R.H.: Routine sampling and the control of *Legionella* spp. in cooling tower water systems. Curr. Microbiol., 2000; 41: 271–275

[8] Burstein D., Zusman T., Degtyar E., Viner R., Segal G., Pupko T.: Genome-scale identification of *Legionella pneumophila* effectors using a machine learning approach. PLoS Pathog., 2009; 5: e1000508
Kubori T., Hyakutake A., Nagai H.: \textit{Legionella pneumophila} translocates an E3 ubiquitin ligase that has multiple U-boxes with distinct functions. Mol. Microbiol., 2008; 67: 1307–1319

Laguna R.K., Creasey E.A., Li Z., Valtz N., Isberg R.R.: A \textit{Legionella pneumophila}-translocated substrate that is required for growth within macrophages and protection from host cell death. Proc. Natl. Acad. Sci. USA, 2006; 103: 18745–18750

Luo L., Luo Z.Q.: The \textit{Legionella pneumophila} effector SidJ is required for efficient recruitment of endoplasmic reticulum proteins to the bacterial phagosome. Infect. Immun., 2007; 75: 592–603

Lu J., Frost L.S.: Mutations in the C-terminal region of TraM provide evidence for in vivo TraM-TraD interactions during F-plasmid conjugation. J. Bacteriol., 2005; 187: 4767–4773

Luedicke C., Slickers P., Ehrlich R., Monecke S.: Molecular finger-printing of \textit{Staphylococcus aureus} from bone and joint infections. Eur. J. Clin. Microbiol. Infect. Dis., 2010; 29: 457–463

Luo Z.Q., Isberg R.R.: Multiple substrates of the \textit{Legionella pneumophila} Dot/Icm system identified by interbacterial protein transfer. Proc. Natl. Acad. Sci. USA, 2004; 101: 841–846

Matuszewski R., Krogiuska B.: Problem występowania palczek \textit{Legionella} w instalacjach i urządzeniach wytwarzających aerosol wodno-powietrzny w obiektach służby zdrowia w Polsce. Nowa Medycyna, 2009; 1: 56–60

Molmerky S.A., Swanson M.S.: \textit{Legionella pneumophila} CurA is a putative repressor of transmission traits and activator of replication. Mol. Microbiol., 2003; 50: 445–461

Monecke S., Ehrlich R.: Rapid genotyping of methicillin-resistant \textit{Staphylococcus aureus}: MRSA isolates using miniaturised oligonucleotide arrays. Clin. Microbiol. Infect., 2005; 11: 825–833

Muller R., Ditzen A., Hille K., Stichling M., Ehrlich R., Illmer T., Ehninger G., Rohayem J.: Detection of herpewirus and adenowirus co-infections with diagnostic DNA microarrays. J. Virol. Methods, 2008; 155: 161–166

Murata T., Delpatio A., Impagnido A., Toomre D.K., Lambright D.G., Roy C.R.: The \textit{Legionella pneumophila} effector protein DrsA is a Rab1 nucleotide-exchange factor. Nat. Cell. Biol., 2006; 8: 971–977

Nagai H., Kagan J.C., Zhu X., Kahn R.A., Roy C.R.: A bacterial guanine nucleotide exchange factor activates ARF on legionella phagosomes. Science, 2002; 295: 679–682

Nederbragt A.J., Balasingham A., Sirevag R., Utkilen H., Jakobsen K.S., Anderson-Glenna M.J.: Multiple-locus variable-number tandem repeat analysis of \textit{Legionella pneumophila} using multi-colored capillary electrophoresis. J. Microbiol. Methods, 2008; 73: 111–117

Ninio S., Celli J., Roy C.R.: A \textit{Legionella pneumophila} effector protein encoded in a region of genomic plasticity binds to Dot/Icm-modified vacuoles. PLoS Pathog., 2009; 5: e1000278

Ninio S., Zuckman-Cholon D.M., Camborone E.D., Roy C.R.: The \textit{Legionella IcmS-IcmW} protein complex is important for Dot/Icm-mediated protein translocation. Mol. Microbiol., 2005; 55: 912–926

Pan X., Lührmann A., Satoh A., Laskowski-Arce M.A., Roy C.R.: Ankyrin repeat repeat analysis of \textit{Legionella pneumophila} serogroup 1 offers the potential for true portability in legionellosis outbreak investigation. J. Clin. Microbiol., 2003; 41: 2932–2939

Gall A., Hoffmann B., Harder T., Grund C., Ehrlich R., Beer M.J.: Rapid haemagglutinin subtyping and pathotyping of avian influenza viruses by a DNA microarray. J. Virol. Methods, 2009; 160: 200–205

Geue L., Schares S., Mintel B., Conraths F.J., Müller E., Ehricht R., Beer M.J.: \textit{Legionella pneumophila} type IV secretion substrates via interdomain horizontal gene transfer. Mol. Microbiol., 2005; 56: 918–933

Buchrieser C.: Multigenome analysis identifies a worldwide distribution of the accidental pathogen \textit{Legionella pneumophila}. J. Clin. Microbiol., 2005; 43: 2047–2052

D’Auria G., Jiménez-Hernández N., Peris-Bondia F., Moya A., Latorre A.: \textit{Legionella pneumophila} pangenome reveals strain-specific virulence factors. BMC Genomics, 2008; 9: 158–166

Ragull S., Garcia-Nuñez M., Pedro-Botet M.L., Sopena N., Esteve M., Segal G., Qu X., Rzhetsky A., Zhang P., Cayanis E., De Jong P.J., Juozaitis V., Pampou S., Georghiou A., Chou I.C., Iannuccilli W., Ulz M.E., Kim D., Asamani G., Hill K., Nuara J., Feder M., Rineer J., Greenberg J.J., Zajac M. et al.: Development of a miniaturized DNA microarray for identification of \textit{Legionella pneumophila} from bone and joint infections. Eur. J. Clin. Microbiol., 2005; 24: 538–544

Sack R.A., Reyes J.E.: The epidemiology of \textit{Legionnaires’ Disease}. EWGLI 2005 European Guidelines for Control and Prevention of Travel Associated Cytomegalovirus. J. Clin. Invest., 1980; 66: 441–450

Joseph C., Lee J., Wijngaarden J., Drasar V., Pastoris, M.C.: \textit{Legionella pneumophila} translocates an E3 ubiquitin ligase that has multiple U-boxes with distinct functions. Mol. Microbiol., 2008; 67: 1307–1319

Habyarimana F., Al-Khodor S., Kalia A., Graham J.E., Price C.T., Geue L., Schares S., Mintel B., Conraths F.J., Müller E., Ehricht R., Beer M.J.: \textit{Legionella pneumophila} type IV secretion substrates interfere with organelle trafficking. PLoS Pathog., 2008; 4: 1000117

Conover G.M., Celli J., Roy C.R.: A \textit{Legionella pneumophila} effector protein DrrA is a Rab1 guanine nucleotide-exchange factor. Nat. Cell. Biol., 2006; 8: 971–977

Segal G., Qu X., Rzhetsky A., Zhang P., Cayanis E., De Jong P.J., Juozaitis V., Pampou S., Georghiou A., Chou I.C., Iannuccilli W., Ulz M.E., Kim D., Asamani G., Hill K., Nuara J., Feder M., Rineer J., Greenberg J.J., Zajac M. et al.: Development of a miniaturized DNA microarray for identification of \textit{Legionella pneumophila} from bone and joint infections. Eur. J. Clin. Microbiol., 2005; 24: 538–544

Shohdy N., Efe J.A., Emr S.D., Shuman H.A.: Pathogen effector protein screening in yeast identifies \textit{Legionella} factors that interfere with membrane trafficking. Proc. Natl. Acad. Sci. USA, 2005; 102: 4866–4871

Suchomitsa R., Farelli J.D., Montminy T.P., Matthews M., Ménètre J.F., Duménil G., Roy C.R., Head J.F., Isberg R.R., Akey C.W.: Structure and function of interacting IcmS-IcmQ domains from a Type IV secretion system in \textit{Legionella pneumophila}. Structure, 2009; 17: 590–601
Viswanathan V.K., Pedersen L.L., Abu Kwaik Y., Krcmarik K., Mody S., Cianciotto N.P.: The cytochrome c maturation locus of \textit{Legionella pneumophila} promotes iron assimilation and intracellular infection and contains a strain-specific insertion sequence element. Infect. Immun., 2002; 70: 1842–1852

Wolfgang M., van Putten J.P., Hayes S.F., Dorward D., Koomey M.: Components and dynamics of fiber formation define a ubiquitous biogenesis pathway for bacterial pili. EMBO J., 2000; 19: 6408–6418

Yu V.L., Plouffe J.F., Pastoris M.C., Stout J.E., Schousboe M., Widmer A., Summersgill J., Heath C.M., Paterson D.L., Chereshky A.: Distribution of \textit{Legionella} species and serogroups isolated by culture in patients with sporadic community-acquired legionellosis: An international collaborative survey. J. Infect. Dis., 2002; 186: 127–128

Zhou H., Ren H., Zhu B., Kan B., Xu J., Shao Z.: Optimization of pulsed-field gel electrophoresis for \textit{Legionella pneumophila} subtyping. Appl. Environ. Microbiol., 2010; 76: 1334–1340

The authors have no potential conflicts of interest to declare.