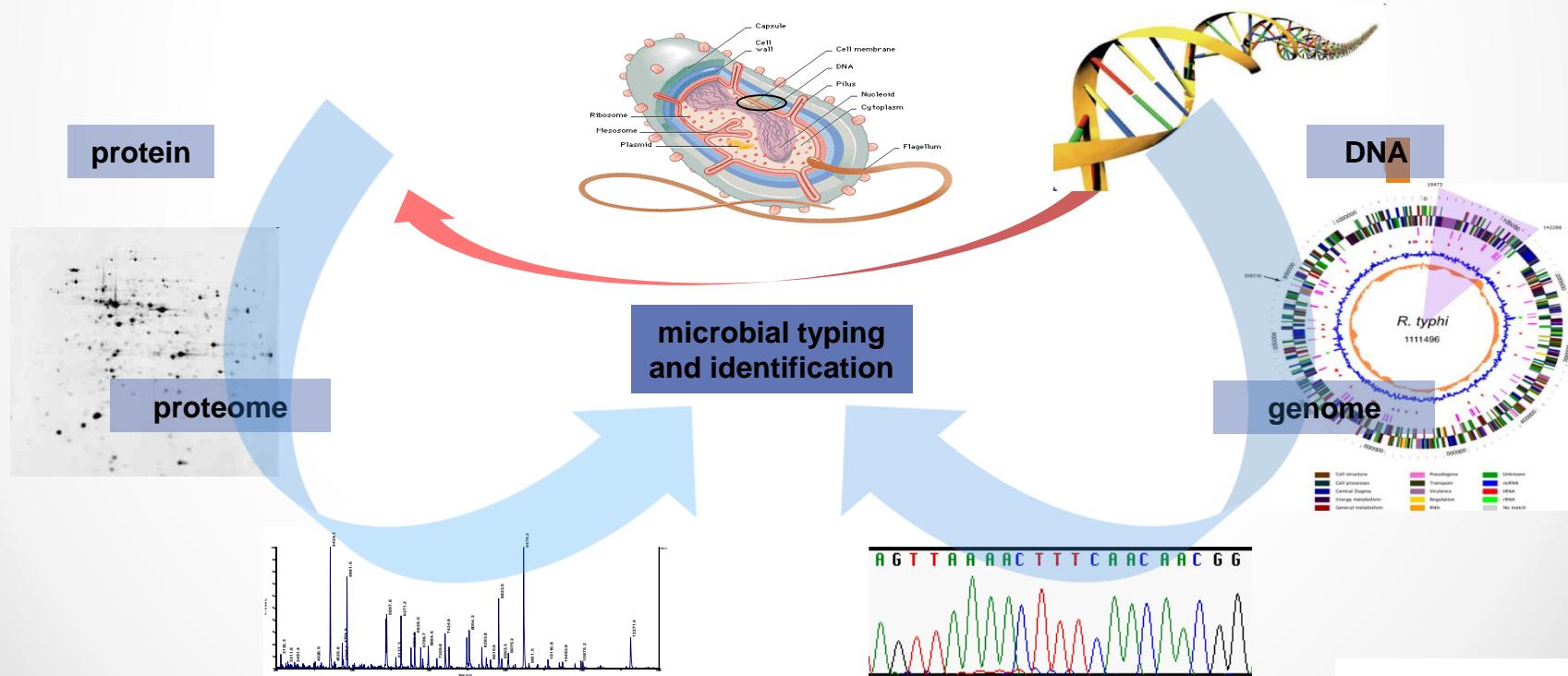


Diagnostics of Infectious Bacteria is Dependent upon a Reliable and Comprehensive Identification and Taxonomy Framework

E. R. B. Moore *et al.*

*CCUG, Department of Clinical Microbiology, Sahlgrenska University Hospital
Department of Infectious Disease, Sahlgrenska Academy, University of Gothenburg, Sweden*



Culture Collection, University of Gothenburg (CCUG)

www.ccug.se

Established 1968 (Curator: E. Falsen, 1968 – 2004; E. Moore, 2004 - ?)

Elisabeth Inganäs, BMA (1991), Maria Ohlén, BMA (1992),
Susanne Jensie-Markopolous, BMA (2002), Sofia Cardew, BMA (2007)
Kent Molin, Chemist (2001), Liselott Svensson-Stadler, Enhetschef (2012)

Created as a unit within the Bacteriological Laboratory of the
Dept Clinical Bacteriology, Sahlgrenska University-Hospital (SU)
SU is a teaching and research hospital

Associated with Univ Gothenburg through Sahlgrenska Academy

Repository for reference strains (> 40,000) of microorganisms and
typing lab for the identification of bacteria from clinical samples
from SU and other hospitals and clinics in Scandinavia

1. *Typing and identification: phenotyping and genotyping*
2. *Archiving, storing, distributing bacterial strains (some few yeast, fungi)*
3. *Research: systematics / diagnostics / antibiotic resistance*

? How do we characterise and identify individual stains
in the diversity of prokaryotes, in the diversity of ecosystems ?

SU BaktLab receives approximately 5,000 samples / week
from the Hospital for microbiological processing

"problematic" strains sent from Routine Lab to
CCUG Typing Lab for analyses

Characterization → typing
Identification → species or sub-species
determinations

What are the “problematic” organisms?

- bacterial strains that are not differentiated from similar and closely related bacteria.
- bacterial strains that are “new”, i.e., they do not fit a database profile;

→ in the CCUG, as many as 150 new 'cases' per year!

Coming from clinical samples!!

→ need effective, reliable characterisation and identification!!!!

Genotypic/Phylogenetic emphasis in bacterial systematics

- *The Bergey's Manuals for Systematic Bacteriology* and *The Prokaryotes* based on 16S rRNA gene sequence-based phylogenies.
- 16S rRNA gene sequence determination and analysis the only single test required for the description of all new bacterial species.
- Genomic DNA-DNA similarities have been the genotypic criterion for defining bacterial species

→ "Gold Standard" ?

¹ "... the complete deoxyribonucleic acid (DNA) sequence would be the reference standard to determine phylogeny and ... phylogeny should determine taxonomy."

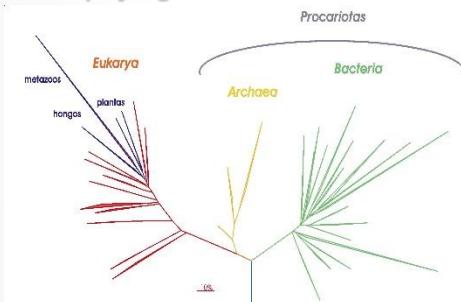
¹ Wayne et al. 1987, Report of the ad-hoc committee on reconciliation of approaches to bacterial systematics, Int J Syst Bacteriol 37: 463-464

Taxonomic Note

Notes on the characterization of prokaryote strains for taxonomic purposes

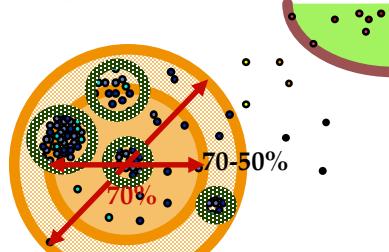
B. J. Tindall,¹ R. Rosselló-Móra,² H.-J. Busse,³ W. Ludwig⁴ and P. Kämpfer⁵

phylogenetic coherence



16S 23S rRNA
Functional genes (MLSA)
Genomic analyses

genotypic coherence



Reassociation DNA-DNA
G+C, PFGE, AFLP, MLSA
Genomic comparisons
(ANI; AAI)

phenotypic coherence



metabolism
chemotaxonomy
proteomics
(MALDI-TOF; LC-MS/MS)

Species descriptions: the examination should be as exhaustive as possible



Genome sequences as the type material for taxonomic descriptions of prokaryotes

William B. Whitman

Show more

<https://doi.org/10.1016/j.syapm.2015.02.003>

[Get rights and content](#)

At recent ICSP Meeting in Valencia, July 2017, this 'proposal' was discussed

Abstract

Genome sequencing of type strains promises to revolutionize prokaryotic systematics by greatly improving the identification of species, elucidating the functional properties of taxonomic groups, and resolving many of the ambiguities in the phylogeny of the higher taxa. Genome sequences could also serve as the type material for naming prokaryotic taxa, which will greatly expand the nomenclature governed by the Bacteriological Code to include many fastidious and uncultured organisms and endosymbionts of great biological interest.

Keywords

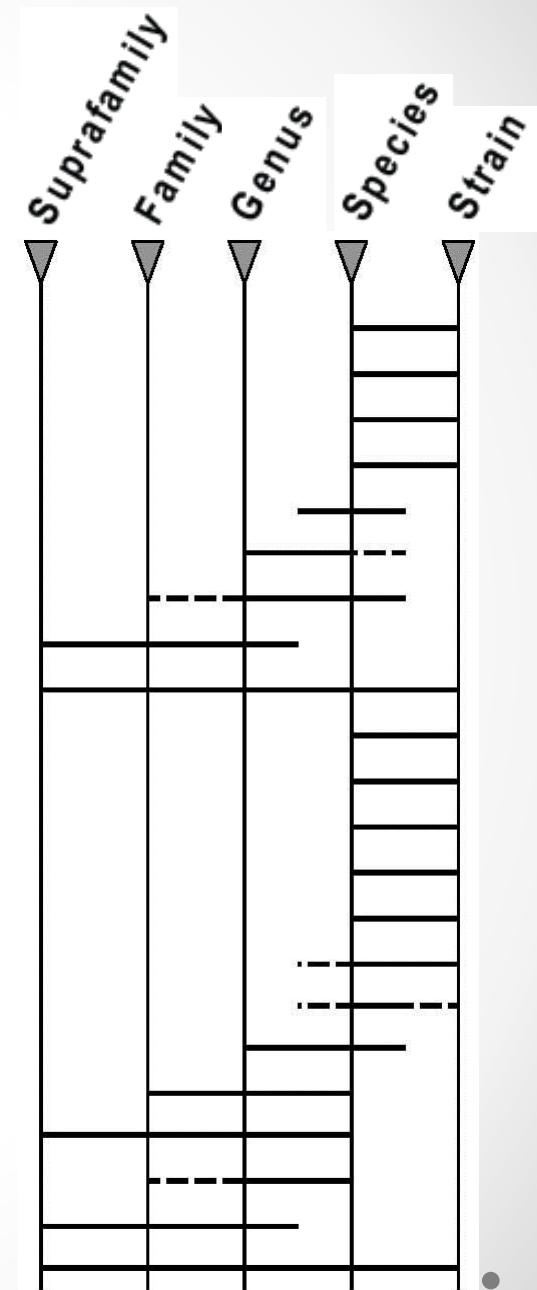
Genomics; Type strain; Bacteriological Code

Bacterial Characterisation and Identification

phenotypic

Methodology

- serotyping (monoclonal and polyclonal antisera)
- Raman spectroscopy
- protein SDS-PAGE (whole-cell or cell envelope)
- pyrolysis mass spectrometry
- Fourier-Transformation Infra-Red (FTIR) spectroscopy
- MALDI-TOF mass spectrometry
- cellular fatty acid fingerprinting (FAME)
- chemotaxonomy (polyamine, polar lipid, quinone, etc., analysis)
- phenotyping (growth, morphology, Api, BIOLOG, etc.)
- restriction analysis of amplified fragments (PCR-RFLP)
- low-frequency restriction fragment analysis (PFGE)
- PCR-amplification analysis (REP-, BOX-, ERIC-PCR, RAPD)
- selective amplification of restriction fragments (AFLP)
- Multi-Locus Variable-number tandem-repeat Analysis (MLVA)
- Ribotyping
- Mulit-Locus Sequence Typing (MLST)
- Multi-Locus Sequence Analysis (MLSA)
- Amplified Ribosomal DNA Restriction Analysis (ARDRA)
- rRNA gene sequence analysis
- DNA-DNA Hybridisation (DDH)
- Genomic DNA mol % G+C
- whole genome sequence analysis



CCUG: since 2013, focussed on developing proteomic-genomic methods for:

strain identification

strain sub-typing

strain characterisation

→ diagnostics of infectious disease in clinical samples without cultivation

Characterisation of infectious or clinically-relevant bacteria

- Virulence – a measure of the degree of pathogenicity

virulence factors – effectively enable:

- colonization in the host (attachment)
- evasion of host defense mechanisms
- potential to cause disease

bacterial toxins, cell surface proteins, hydrolytic enzymes, ...

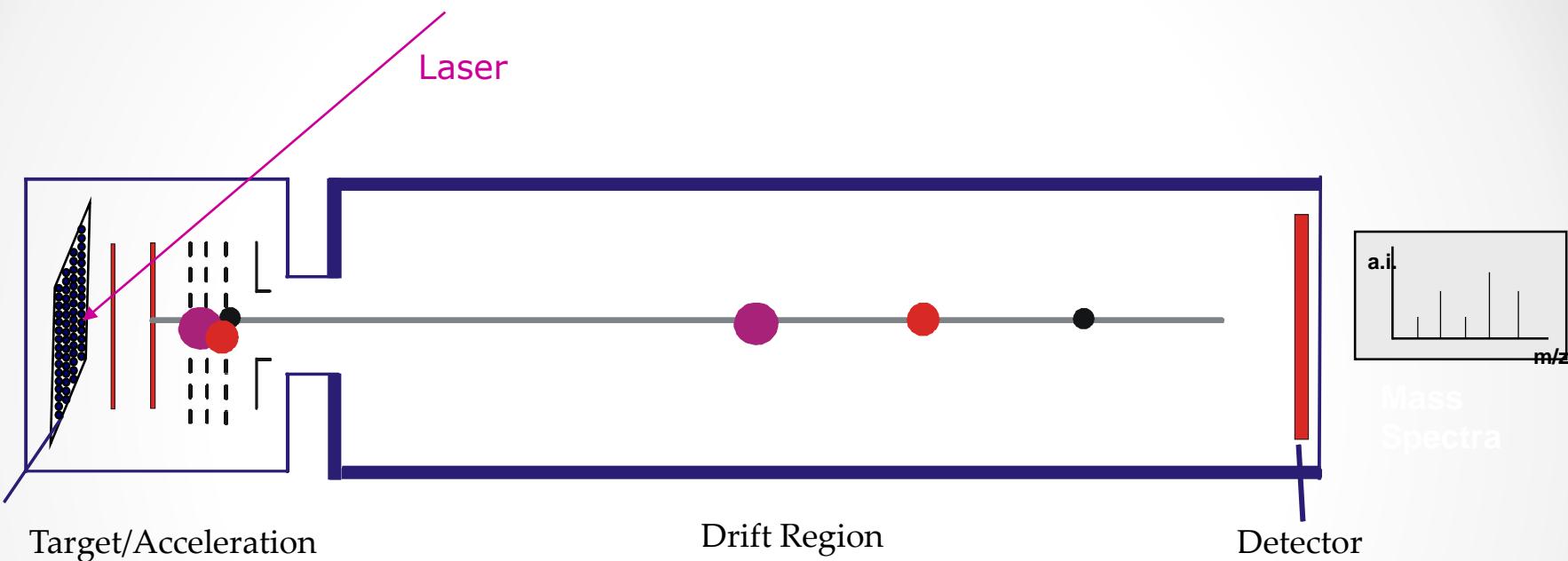
- Antibiotic resistance – inherent and acquired resistance

WHO – "*... one of the biggest threats to global health, food security, and development today. ...we are heading for a ‘post-antibiotic era’, in which common infections and minor injuries can once again kill.*"

resistance factors - enable:

- survival in competitive environments
 - pathogenic bacteria to adapt to antimicrobial therapies
- gene activation/repression, binding proteins, porins, efflux pumps, ...

MALDI-TOF MS



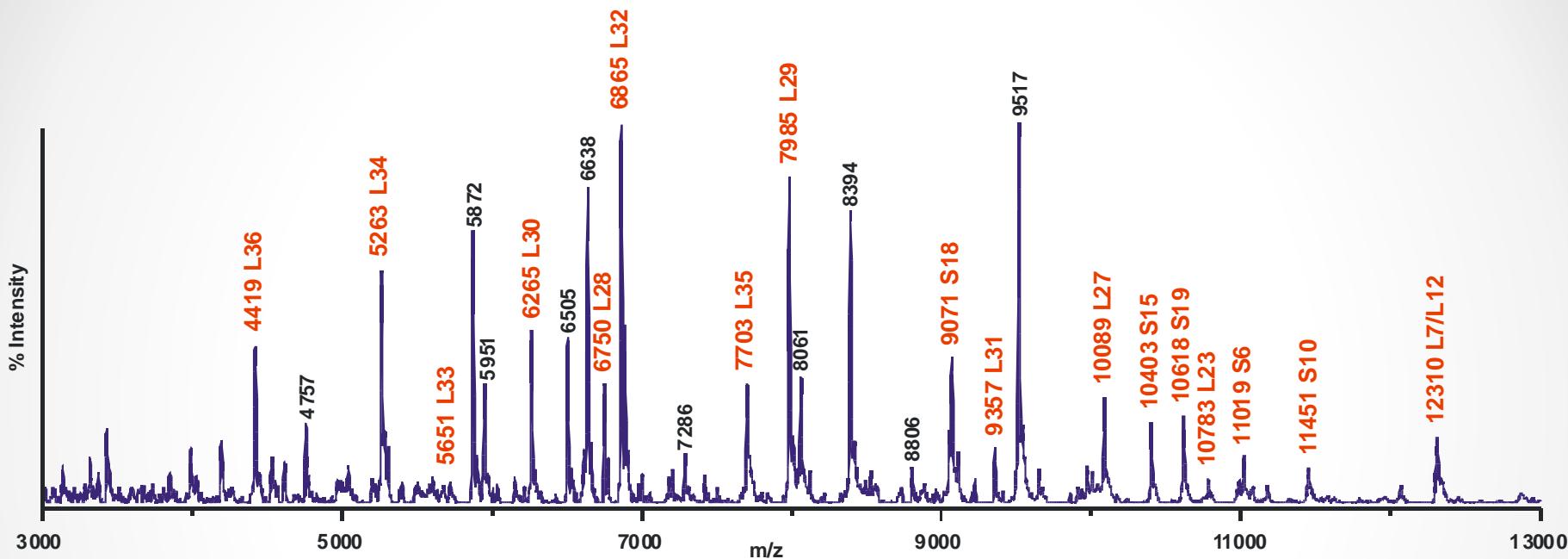
Desorption/Ionisation

Time-of-Flight

Molecular Mass



Streptococcus pneumoniae: identifying mass signals

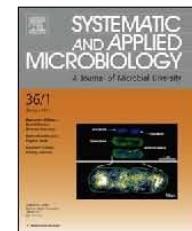
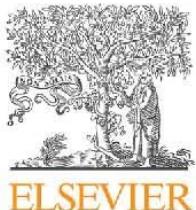


mass signals matching to ribosomal proteins are marked in red

~50% of peaks represent ribosomal proteins

protein masses obtained from the Swiss-Prot/TrEMBL sequence database

"-omics" relevant for identification of proteins



Review

Proteotyping: Proteomic characterization, classification and identification of microorganisms – A prospectus



Roger Karlsson^{a,b,*}, Lucia Gonzales-Siles^c, Fredrik Boulund^d, Liselott Svensson-Stadler^{a,c}, Susann Skovbjerg^{a,c}, Anders Karlsson^b, Max Davidson^b, Stefan Hulth^e, Erik Kristiansson^d, Edward R.B. Moore^{a,c,f}

^a Department of Clinical Microbiology, Sahlgrenska University Hospital, SE-41346 Gothenburg, Sweden

^b Nanoxis Consulting AB, SE-40016 Gothenburg, Sweden

^c Department of Infectious Diseases, Institute of Biomedicine, Sahlgrenska Academy of the University of Gothenburg, SE-40234 Gothenburg, Sweden

^d Department of Mathematical Sciences, Chalmers University of Technology, SE-41296 Gothenburg, Sweden

^e Department of Chemistry and Molecular Biology, University of Gothenburg, SE-41296 Gothenburg, Sweden

^f Culture Collection University of Gothenburg (CCUG), SE-41346 Gothenburg, Sweden

ARTICLE INFO

Keywords:

Proteotyping

Proteomics

Mass spectrometry

Microbial systematics

ABSTRACT

Modern microbial systematics requires a range of methodologies for the comprehensive characterization, classification and identification of microorganisms. While whole-genome sequences provide the ultimate reference for defining microbial phylogeny and taxonomy, selected biomarker-based strategies continue to provide the means for the bulk of microbial systematic studies. Proteomics, the study of the expression of genes, as well as the structure and function of the resulting proteins, offers indirect measures of genome sequence data. Recent developments in applications of proteomics for analyzing

Can tandem Mass Spectrometry and proteomics – 'Proteotyping', be applied for the direct analyses of clinical samples for rapid and reliable diagnostics of infectious bacteria?

What would be the advantages over traditional diagnostic methods or NGS-based methods?

1) Respiratory infections: detection and identification, virulence and antibiotic resistance

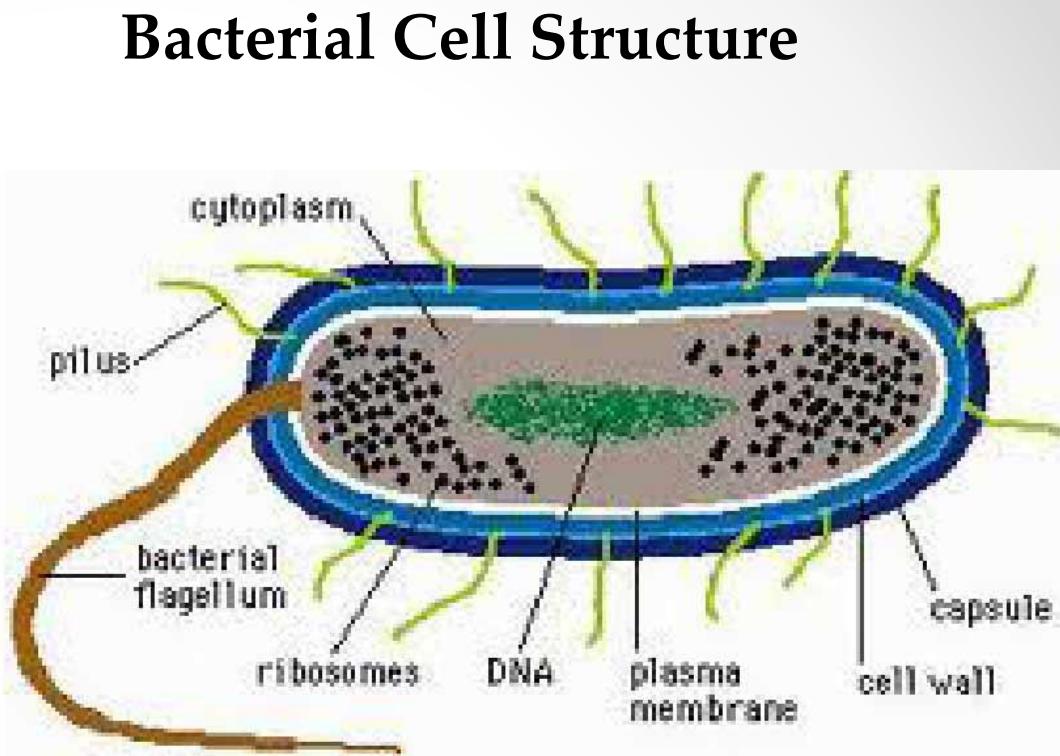
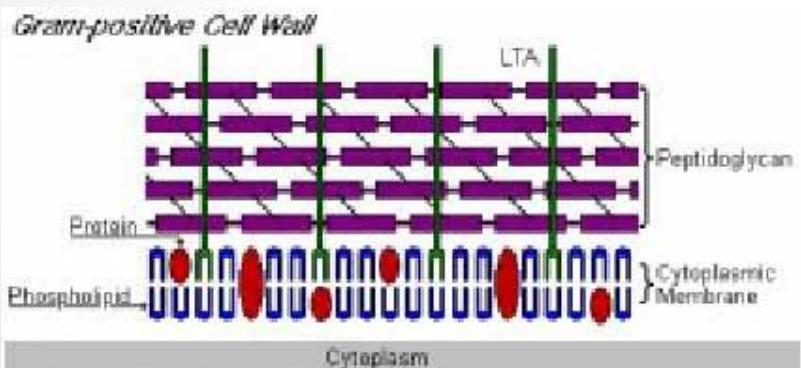
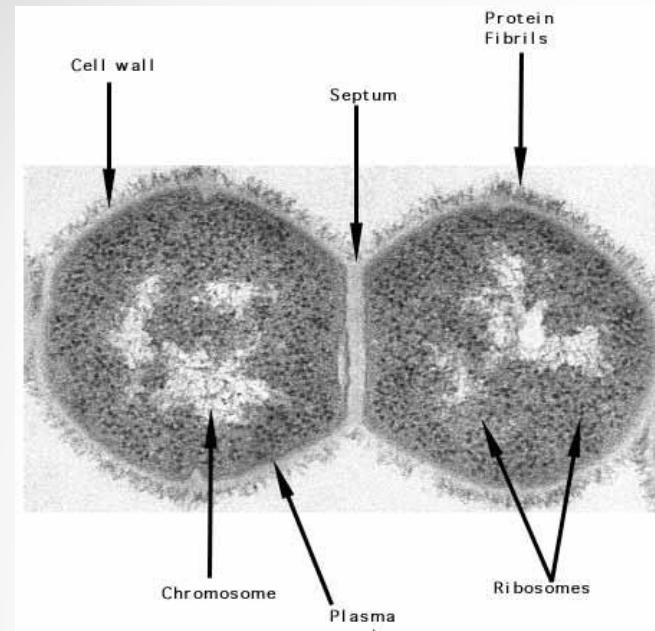
*S. pneumoniae, S. aureus, P. aeruginosa, K. pneumoniae,
S. pyogenes, H. influenzae, M. catarrhalis, M. pneumoniae*

2) Urinary-tract infections: detection and identification, virulence and antibiotic resistance

E. coli, K. pneumoniae, Enterobacteriaceae

3) Septicemia and Sepsis: detection and identification, virulence and antibiotic resistance

E. coli, S. aureus, S. pyogenes



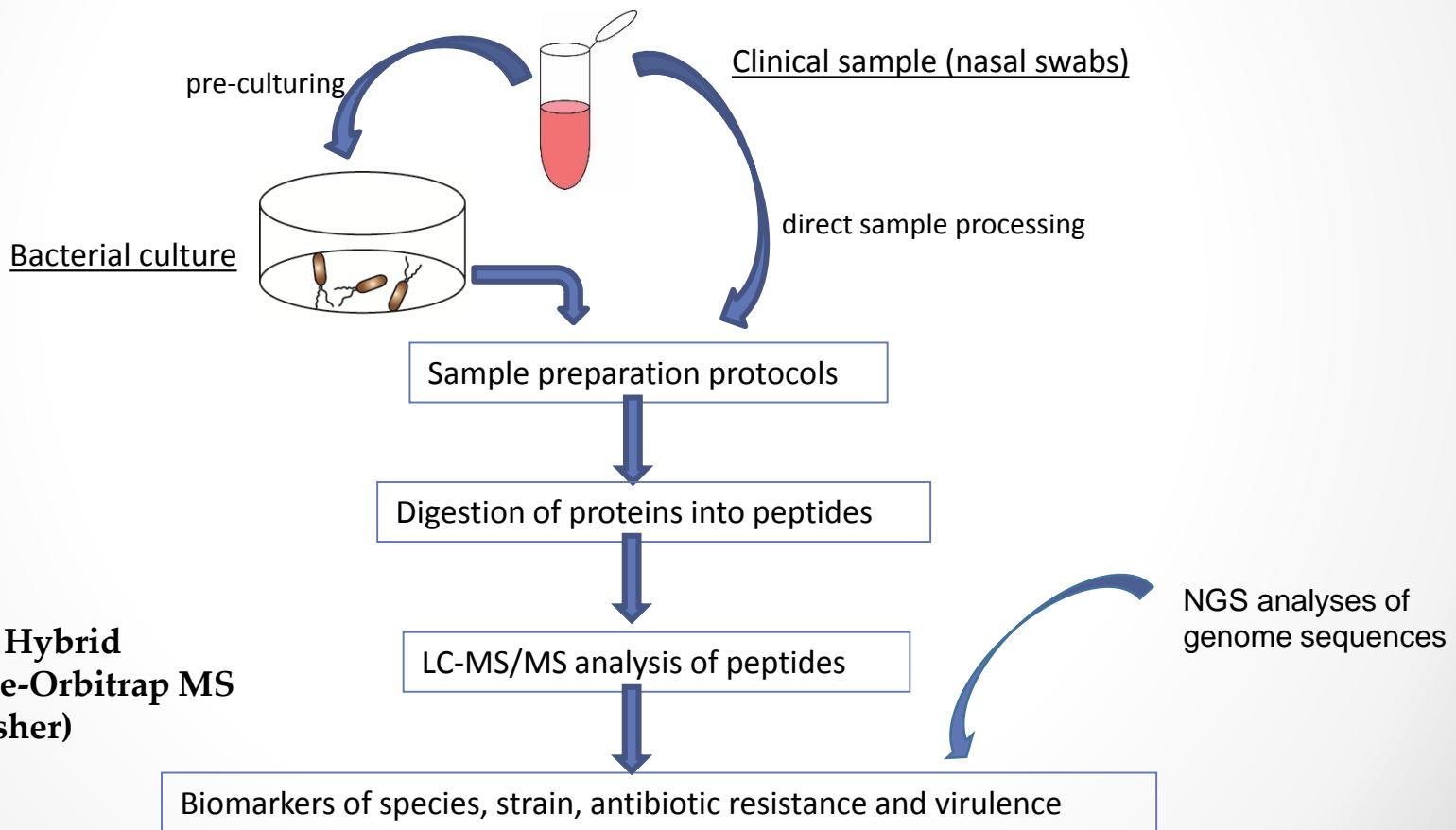
Optimisation of MS-proteomics

Different protein targets in different cell compartments.

* optimised fractionation of Gram-positive and Gram-negative cells.

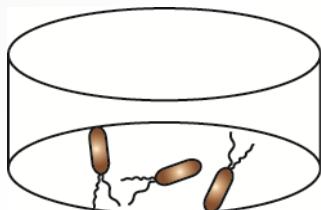
LC-MS/MS 'bottom-up' Proteomics-based bacterial identification and diagnostics

Sample and Workflow

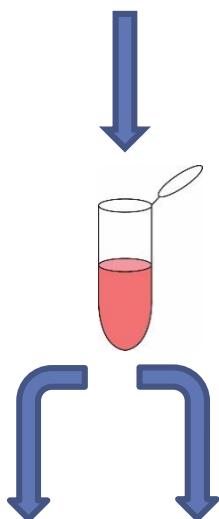


LC-MS/MS-based Proteomics

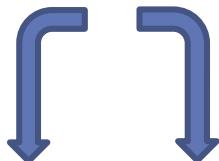
Bacterial culture
Clinical sample



Wash cells



*Bead beating
and centrifugation
Fractions*



Sample preparation strategy

Intact cells - Surface shaving

Surface proteins = hotspots for genetic variation!
Virulence factors – Host-pathogen interactions

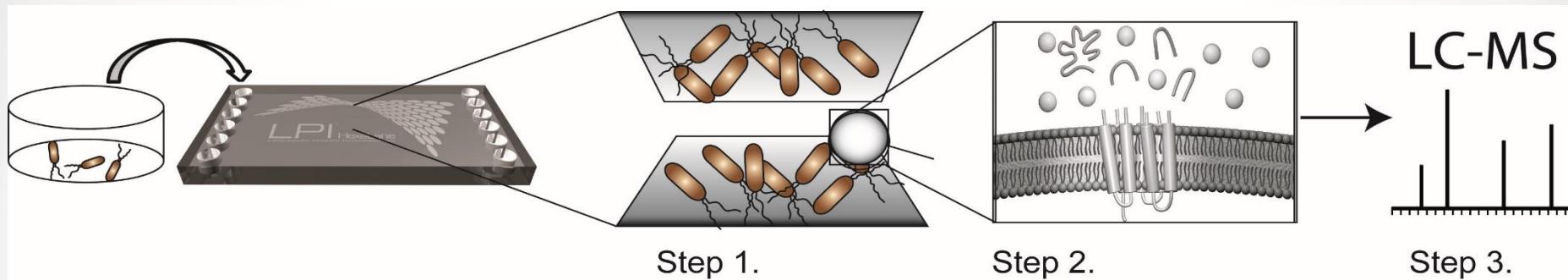
Membrane and soluble fractions

Highly abundant conserved proteins = species markers!
Antibiotic resistance markers and virulence factors

Soluble

Membrane

Peptide generation by LPI FlowCell analysis



Step 1. Microbial cells or cell fractions are immobilized in the Hexalane Flow-Cell.

Step 2. Trypsin (digestive enzyme) is added and allowed to digest exposed proteins

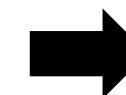
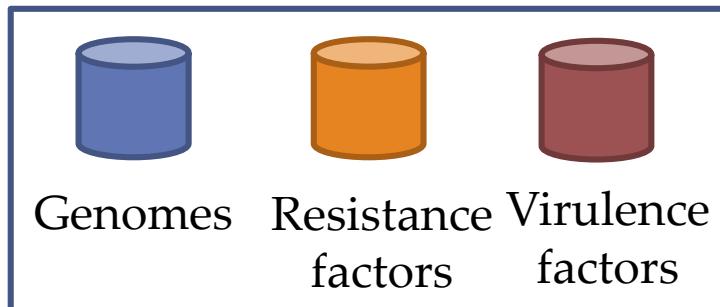
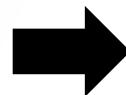
Step 3. Peptides are eluted from the Flow-Cell and analyzed, using LC-MS/MS

Trypsin digestion: carboxyl-side of lysine or arginine, except if followed by proline

Shotgun 'bottom-up' proteomics – data processing

Peptide fragments

LIEVDSQGNR
TVFFEDGNQPISK
LLEDVGSGNGAR
ILEVNEER
LLDIDQEER
AGVGEAAGRGAAGVGGR



Proteome profile



Sample statistics on strain level

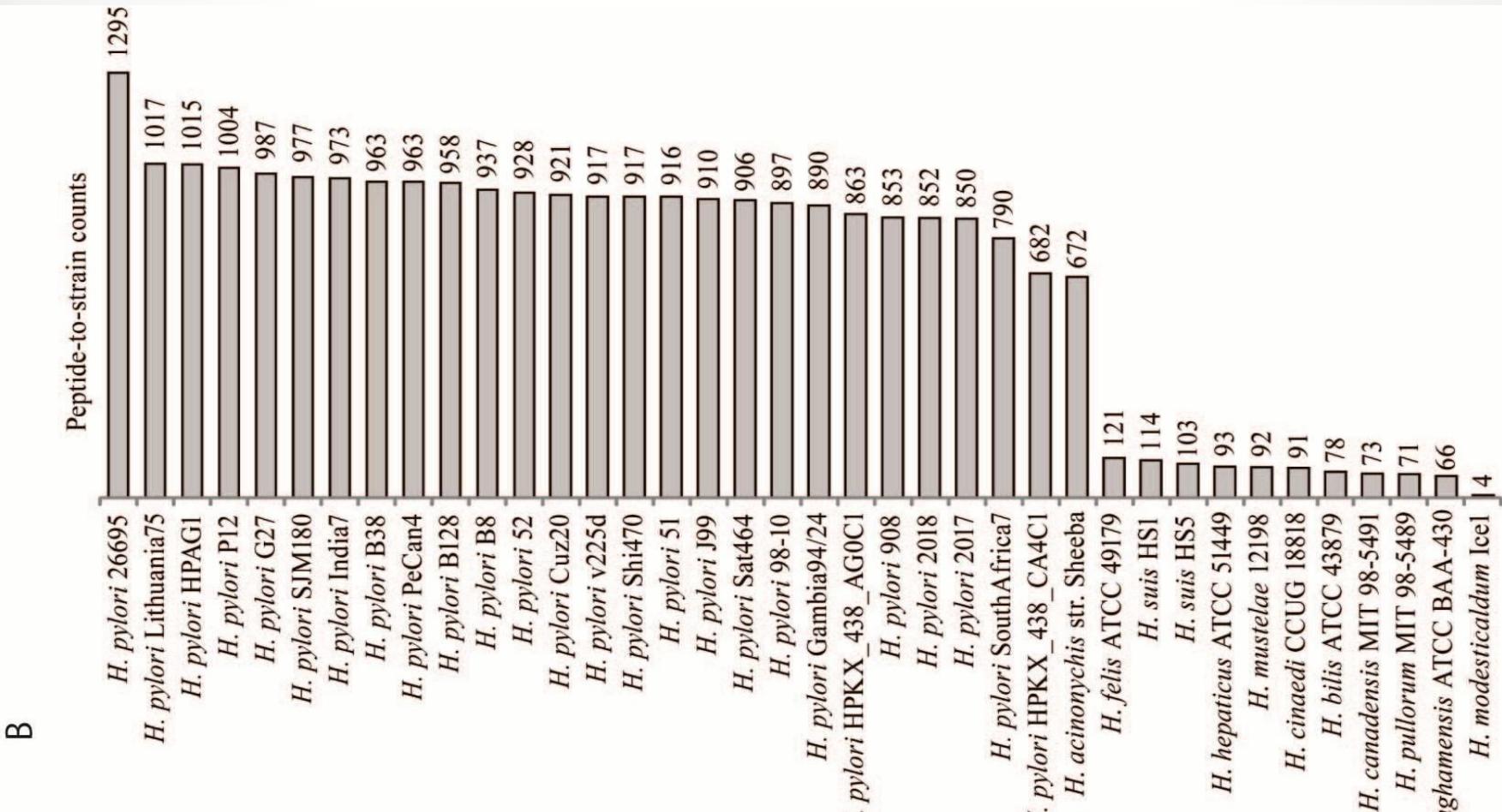
Sensitive matching

- Genome-wide alignment in all six reading-frames
- Average size of matched fragment: 16 AA (down to 6 AA)
- Database
 - well-annotated genomes
 - >5000 resistance factors (E Kristiansson, Chalmers Univ. Technology)
 - Virulence Factor Database: www.mgc.ac.cn/VFs/main.htm

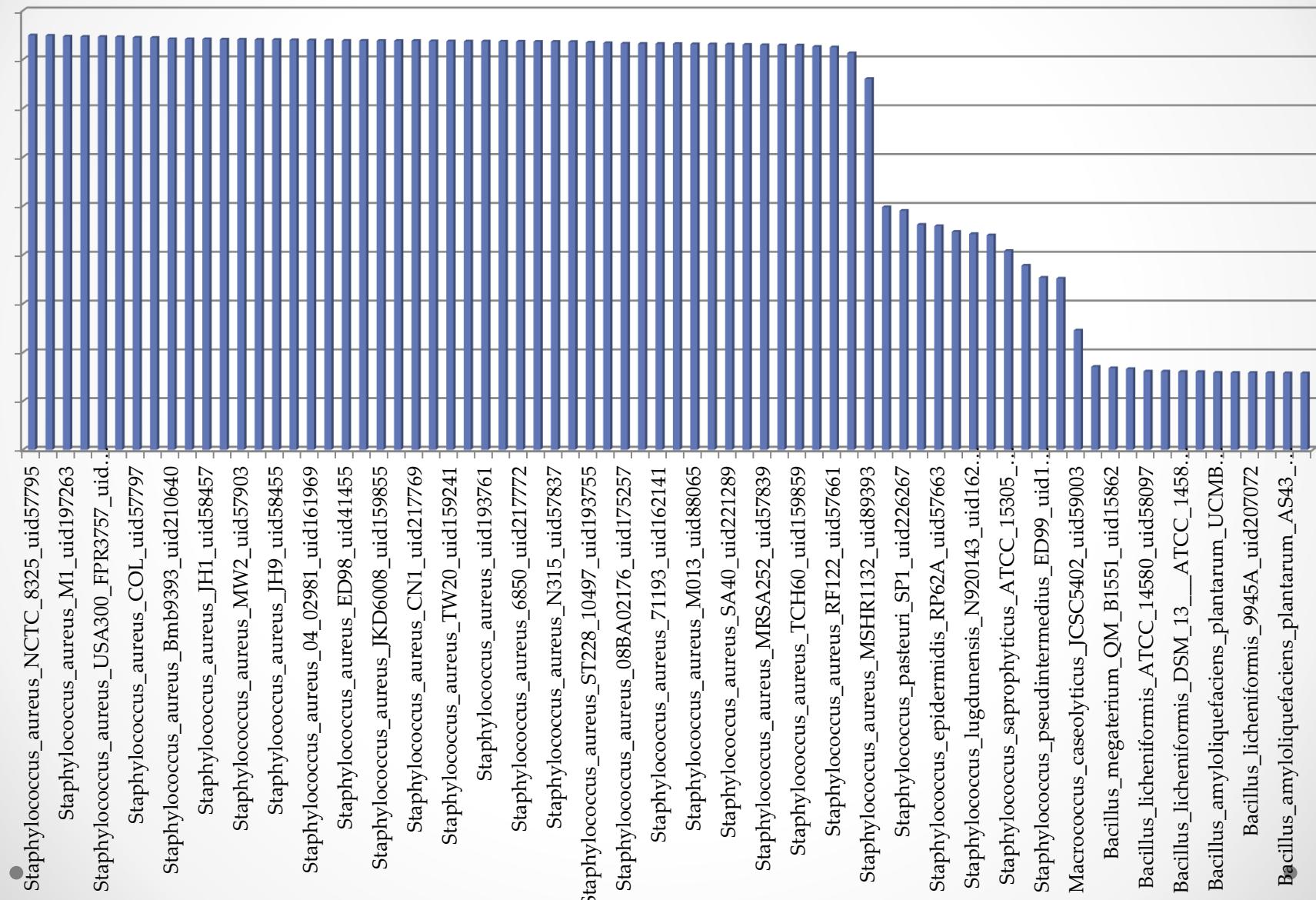
→ Peptide biomarkers

Proteotyping: *Helicobacter pylori* – ranking species matching

Analysis of the *H. pylori* strain 26695 sample. The distribution of peptide-to-strain counts for different *H. pylori* strains available in databases is shown.

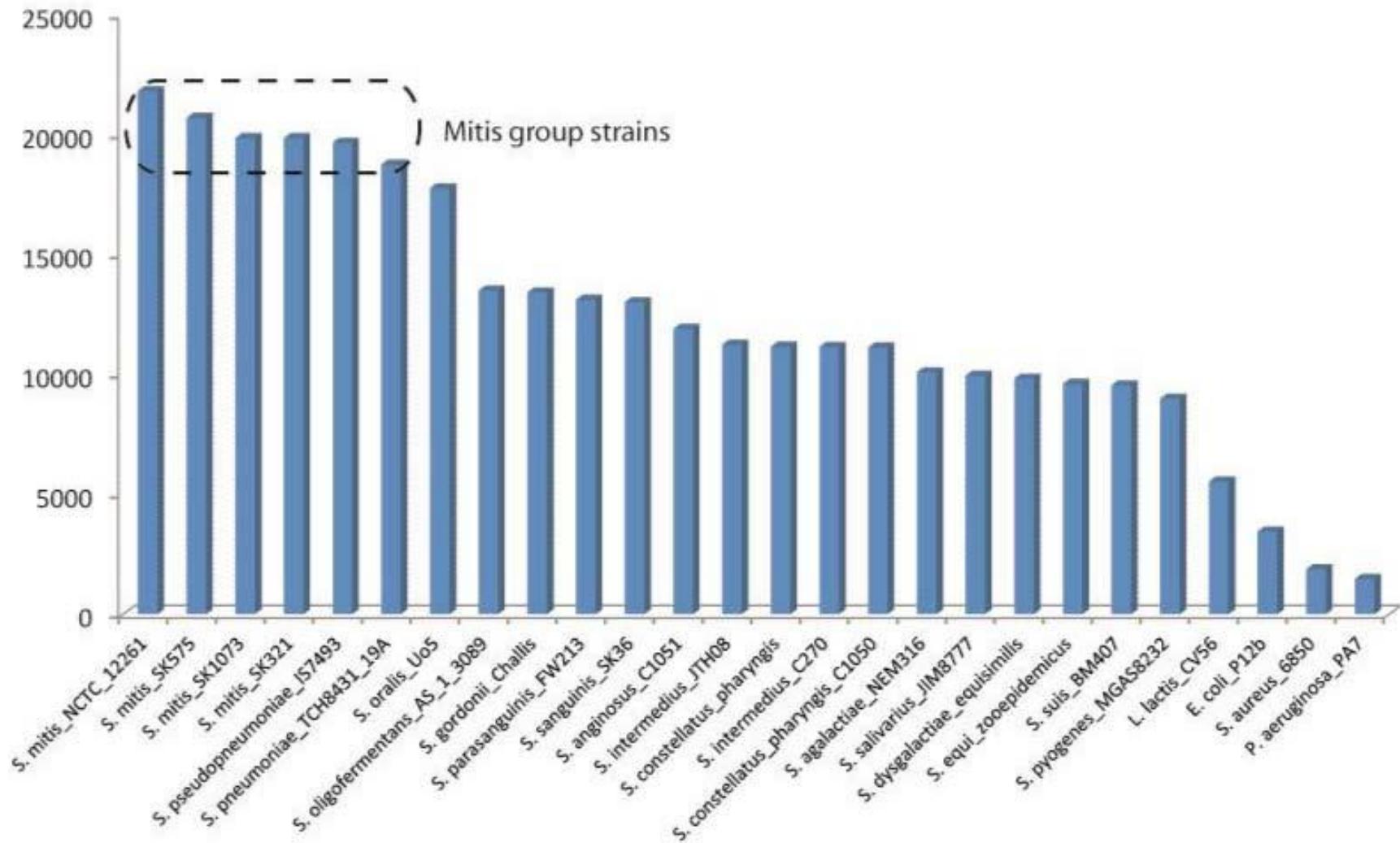


Proteotyping: *Staphylococcus aureus* – ranking species matching



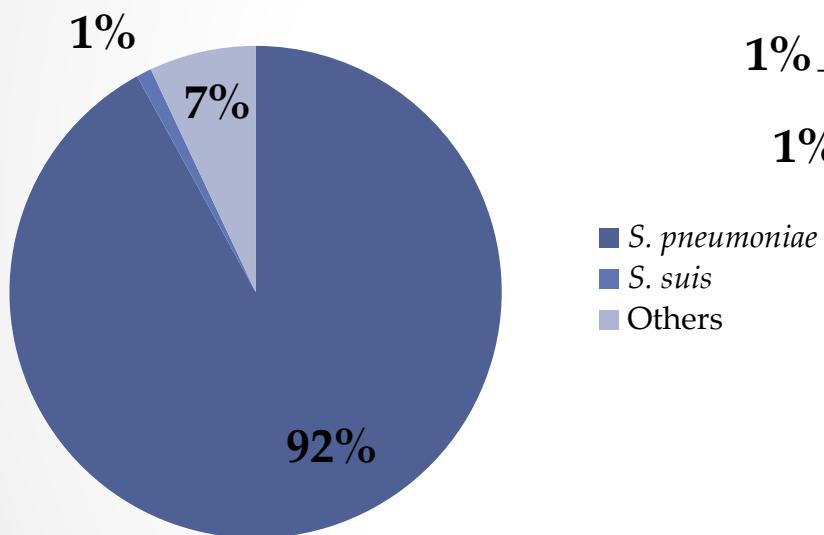
Proteotyping: *Streptococcus* spp. – ranking species matching

S. mitis ranking

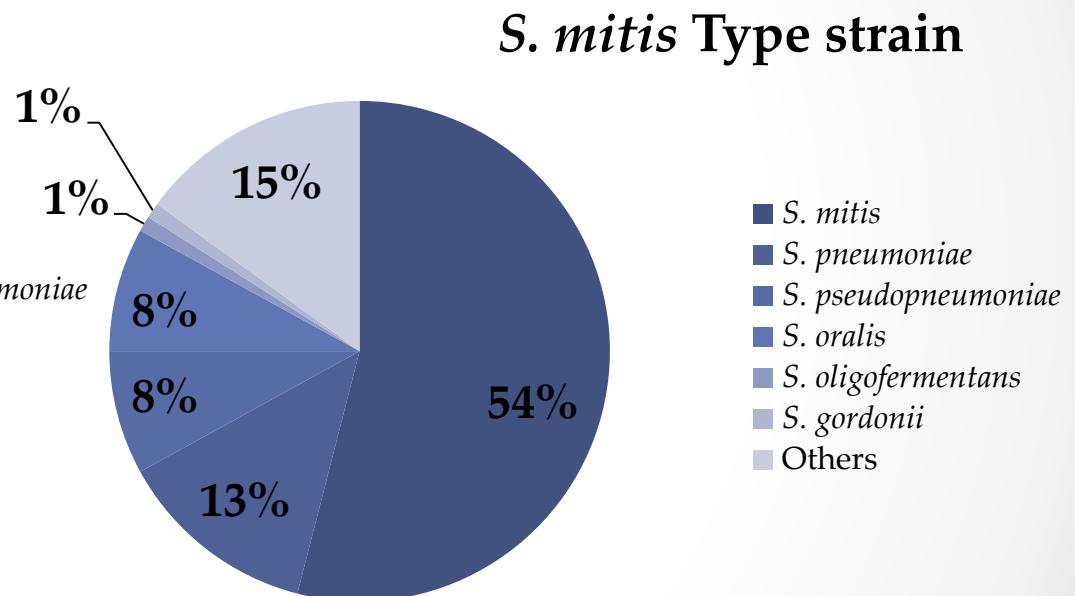


Accuracy in correct species identifications in the Mitis-Group

S. pneumoniae Type strain



S. mitis Type strain



- Only 1 WGS of *S. pseudopneumoniae* and *S. mitis* available in the database
- *S. mitis* Type strain was not available – the B6 strain genome clustered far away from the Type strain

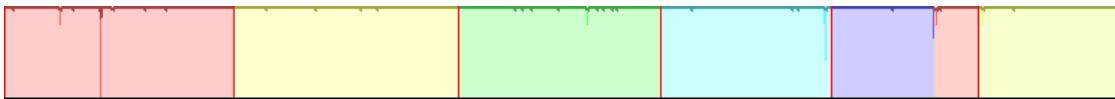
**How can we know if a genome
is correctly classified ?**

ANI: Average Nucleotide Identity

Genome
sequence 1

VS

Genome
sequence 2



Fragments of the genome sequences are aligned and compared



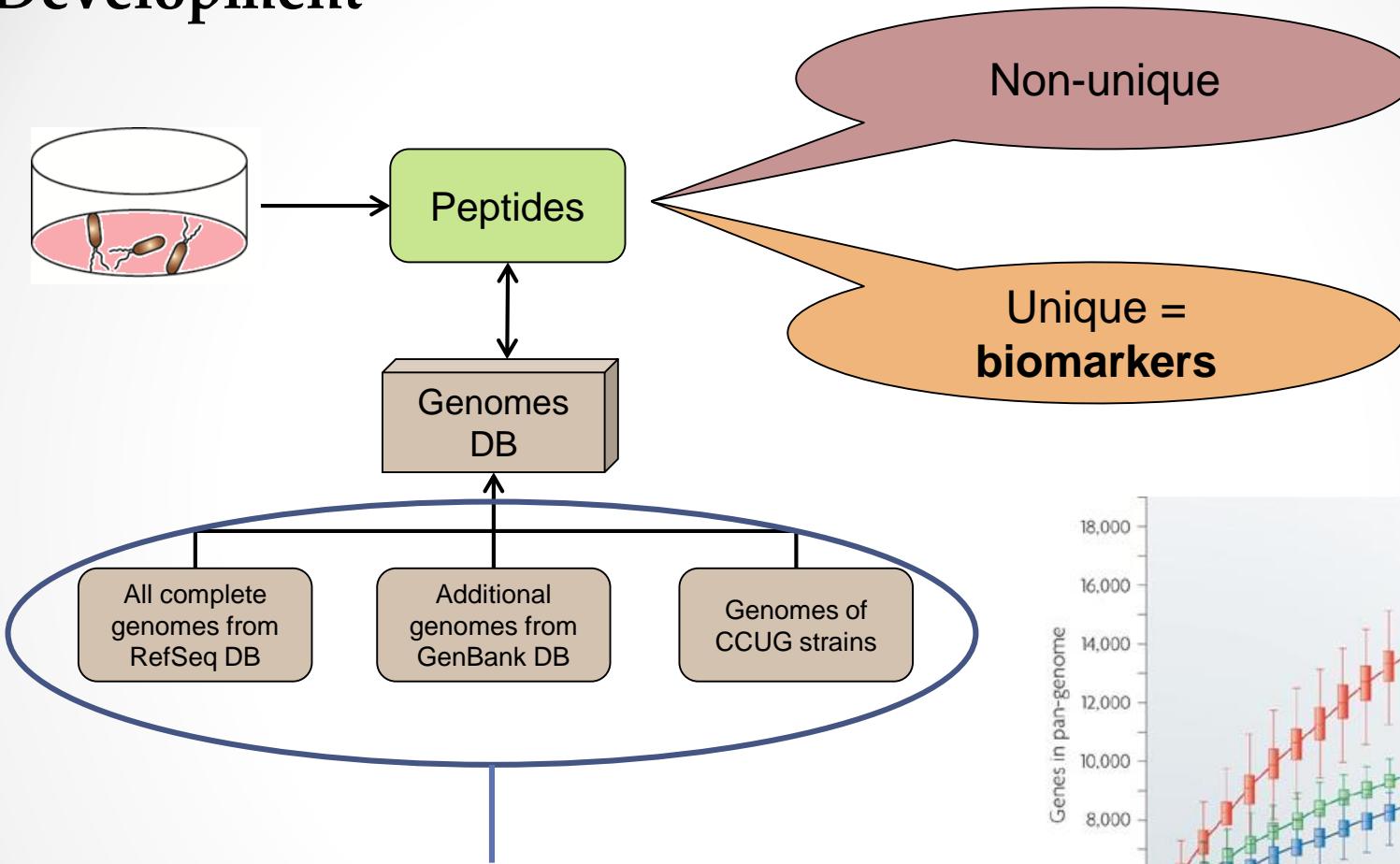
Similarity value

ANI value between
two genomes

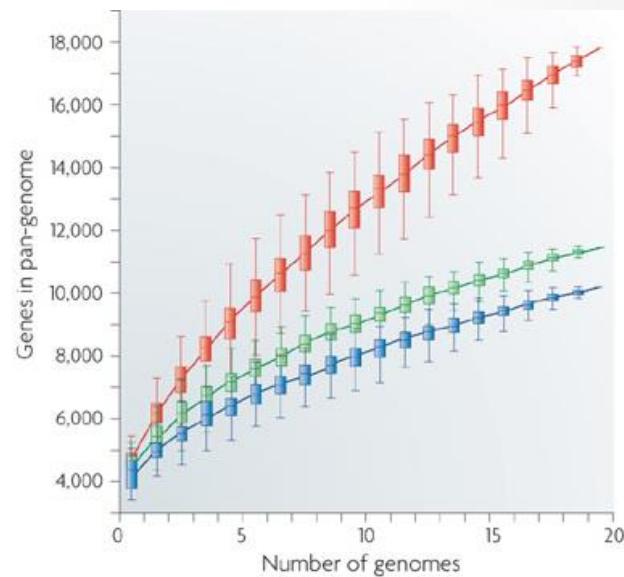
$> 96\%$	= Same species
$93 - 96\%$	= Species boundary
$< 93\%$	= Different species

Organism	Total number of genome sequences	ANIB		
		< 93%	93 - 96%	≥ 96%
<i>Streptococcus pneumoniae</i>	323	0	0	323
<i>Streptococcus mitis</i>	40	18	22	0
<i>Streptococcus australis</i>	0	0	0	0
<i>Streptococcus cristatus</i>	2	0	1	1
<i>Streptococcus infantis</i>	4	3	1	0
<i>Streptococcus oligofermentans</i>	7	7	0	0
<i>Streptococcus oralis</i>	11	7	4	0
<i>Streptococcus parasanguinis</i>	9	1	6	2
<i>Streptococcus peroris</i>	0	0	0	0
<i>Streptococcus pseudopneumoniae</i>	6	0	0	6
<i>Streptococcus sanguinis</i>	23	2	16	5
<i>Streptococcus sinensis</i>	0	0	0	0
<i>Streptococcus tigurinus</i>	3	0	3	0
Total	408	38	53	3B7
	100%	35.88%	50.38%	18.34%

Proteomic-Genomic analyses Development



We need more
genomes!!



Tenaillon et al. 2010

CCUG Genomics

- Started genome sequencing, January 2015
- To-date: approx 200 draft whole-genome sequences



Ion-Torrent
(SU Clin
Microiol)

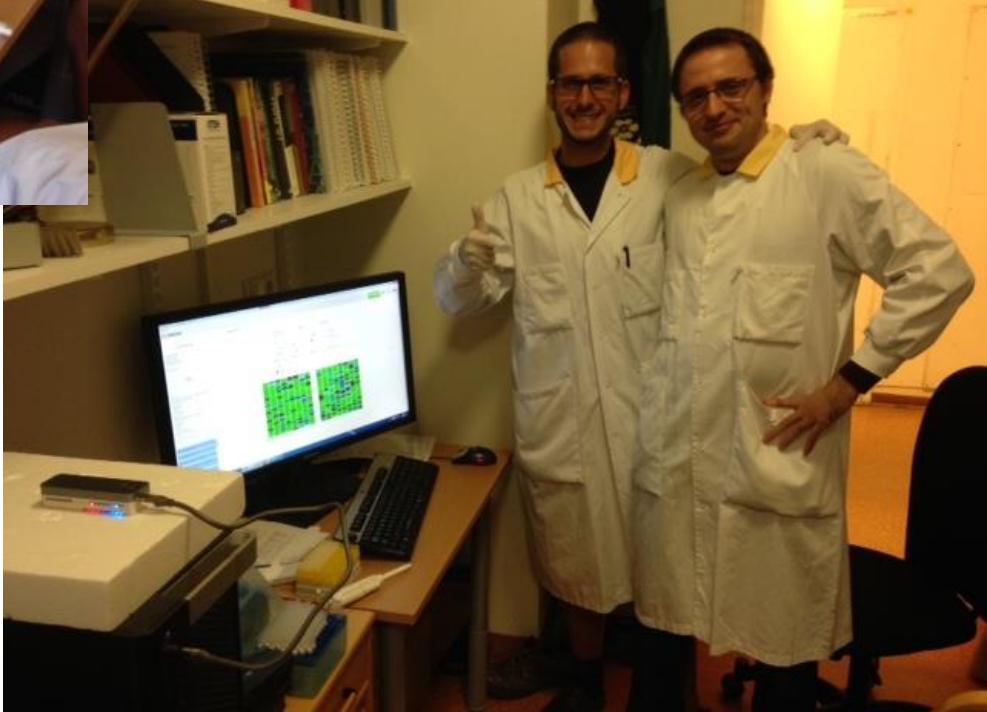
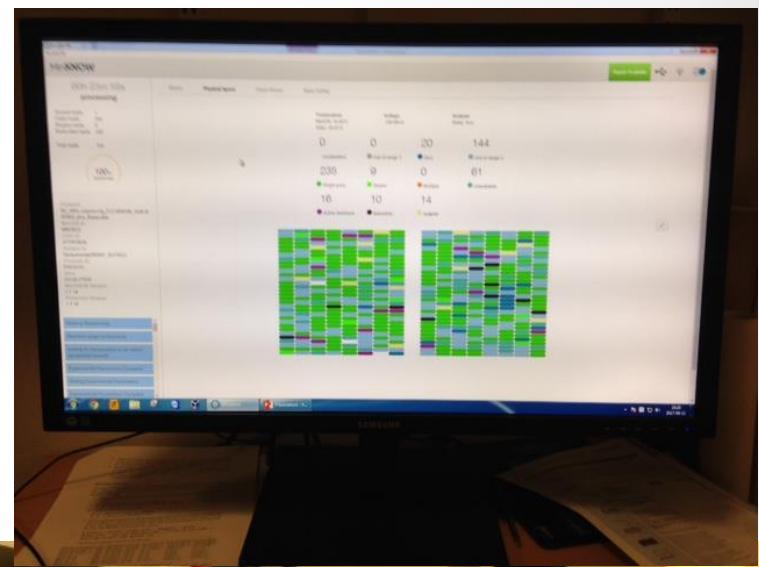
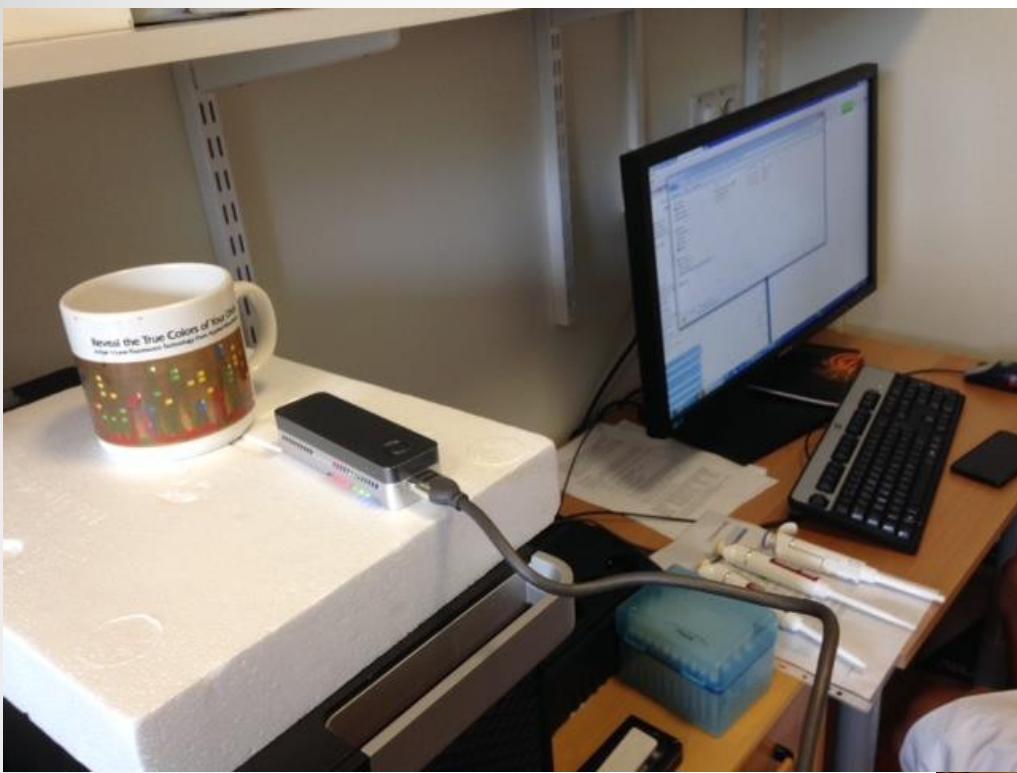


Illumina
(SU Genomics
Core)



PacBio
(NGI – Uppsala)

Introduction of Oxford Nanopore MinION to CCUG

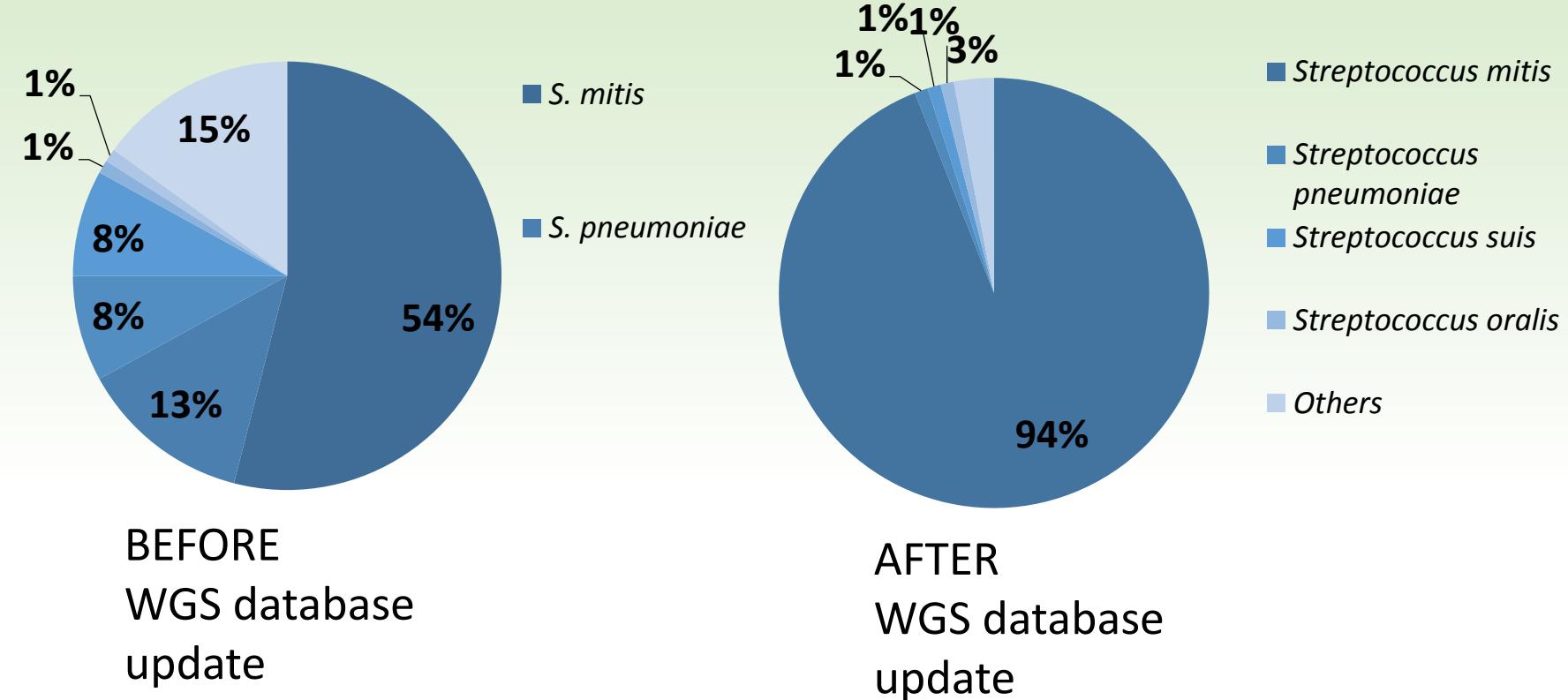


Addition of *Streptococcus* spp. genomes to the database

Organism	2015-02-10	2016-12-31
<i>Streptococcus australis</i>	0	1
<i>Streptococcus cristatus</i>	0	1
<i>Streptococcus gordonii</i>	1	3
<i>Streptococcus infantis</i>	0	1
<i>Streptococcus mitis</i>	1	33
<i>Streptococcus oligofermentans</i>	1	1
<i>Streptococcus oralis</i>	1	13
<i>Streptococcus parasanguinis</i>	2	3
<i>Streptococcus peroris</i>	0	1
<i>Streptococcus pneumoniae</i>	27	28
<i>Streptococcus pseudopneumoniae</i>	1	9
<i>Streptococcus sanguinis</i>	1	6
<i>Streptococcus sinensis</i>	0	1
<i>Streptococcus tigurinus</i>	0	4
	35	105

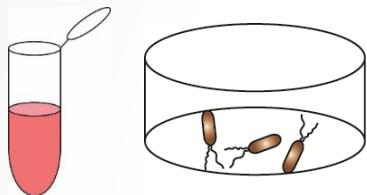
Accuracy in finding the correct species in the S. Mitis-Group

Streptococcus mitis Type strain

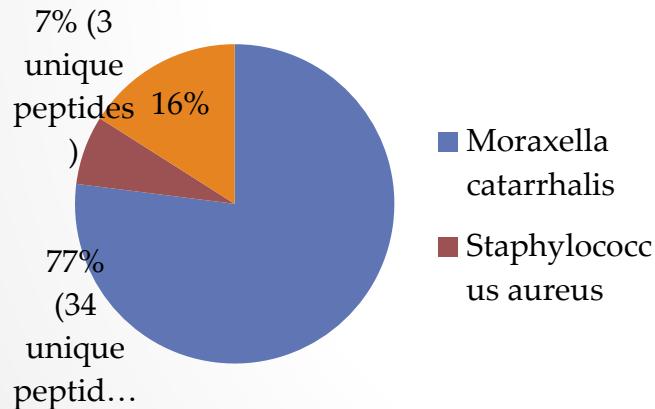


Proteotyping of Sahlgrenska Hospital clinical samples (nasopharyngeal) Positive samples

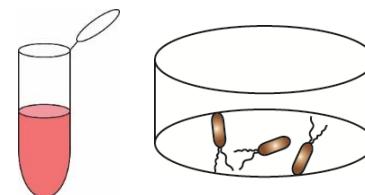
Example 1



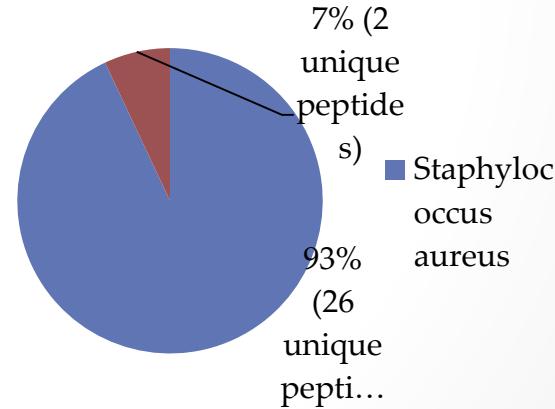
Sample **positive** for
Moraxella catarrhalis



Example 2



Sample **positive** for
Staphylococcus aureus



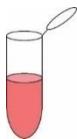
Results of MS-Proteomics peptide identification and matching

Direct MS-proteomics analysis of clinical sample

Biomarker inclusion list for *S. pneumoniae*

Samples from Utrecht Proteomic results

EU516 Clinical sample
From Utrecht Pediatrics
Clinic



UM 6044

High number of bacteria

qPCR positive for
✓ *S. pneumoniae*
✓ *S. aureus*

Taxonomy view from Uniprot	Number of protein hits
Bacteria (152)	
Firmicutes (148)	
Bacilli (144)	
Lactobacillales (49)	
Lactobacillus reuteri 100-23 (1)	
Streptococcus (48)	
Streptococcus pneumoniae (42) ←	
Streptococcus pseudopneumoniae (2)	
Streptococcus sanguinis (4)	
Staphylococcus (95) ←	
Staphylococcus aureus (91)	
Staphylococcus hominis subsp. hominis C80 (1)	
Staphylococcus simiae CCM 7213 (2)	
Staphylococcus sp. M0480 (1)	
Clostridiales (3)	
Lachnospiraceae (2)	
Lachnospiraceae bacterium 8_1_57FAA (1)	
Ruminococcus torques ATCC 27756 (1)	
Syntrophomonas wolfei subsp. wolfei (strain DSM 2245B / Goettingen) (1)	
Mitsuokella sp. oral taxon 131 str. W9106 (1)	
Mycobacterium vanbaalenii (strain DSM 7251 / PYR-1) (1)	
Mycoplasma sp. CAG:472 (1)	
Proteobacteria (2)	
Pseudomonas fulva (strain 12-X) (1)	
Rhodopseudomonas palustris (strain BisB5) (1)	
Eukaryota (309)	

Proteotyping of *S. mitis* complex

All strains analysed in triplicate
Good reproducibility
Identified peptides and proteins
Identified discriminatory peptides

Discriminatory proteins

S. pneumoniae discriminatory proteins found by proteotyping workflow

- general stress protein 24
- surface protein PspA
- rpsP 30S ribosomal protein S16
- strH beta-N-acetylhexosaminidase
- cbpA choline binding protein A
- acpP acyl carrier protein
- ugd UDP-glucose 6-dehydrogenase
- prsA foldase PrsA
- lytB endo-beta-N-acetylglucosaminidase
- **SURFACE-ASSOCIATED VIRULENCE FACTORS**

	<i>S. pneumoniae</i>	<i>S. oralis</i>	<i>S. mitis</i>	<i>S. pseudopneumoniae</i>	<i>S. pyogenes</i>
Number of proteins	649	564	678	622	407
Number of peptides	3556	2495	3722	2141	1197
Number of discriminative peptides	417	282	237	313	308

S. pneumoniae virulence factors

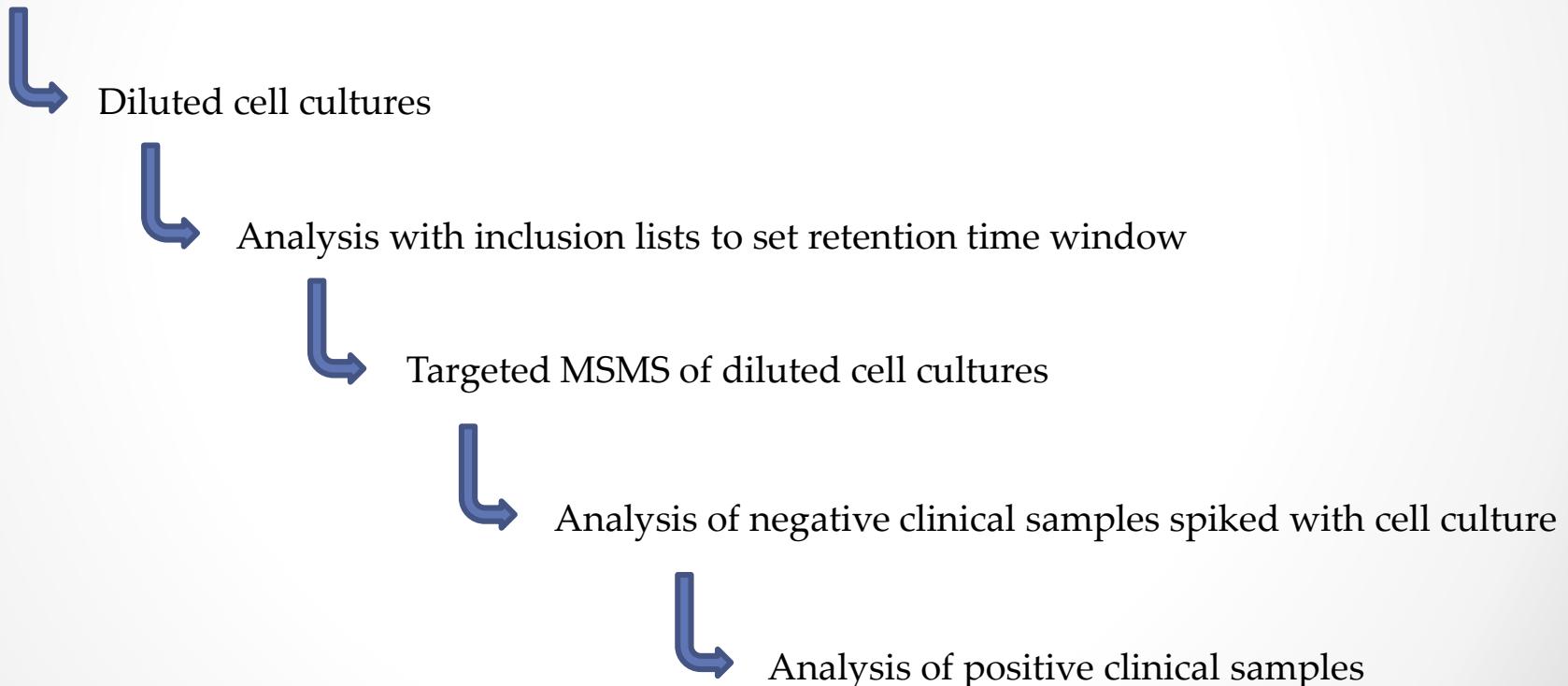
		S. pne	S. pse	S. mit
PspA	Pneumococcal surface protein A	+++	N	N
LytA	Autolysin A	++	N	N
Bac	Bacteriocin	++	++	N
StrH	B-N-acetylglucosaminidase	+++	?	?
BgaA	B-galactosidase	++	?	?
CpbA	Choline binding protein A	+++	N	N
Eno	Enolase	++	+	+
Hyl	Hyaluronate lyase	N	N	N
IgA	IgA1 protease	++	N	N
Nan	Neuraminidase	++	N	N
ChoP	Phosphorylcholine	?	?	?
PavA	Pneumococcal adhesion/virulence A	++	N	N
PlaA	Pneumococcal iron acquisition A	?	N	N
PiuA	Pneumococcal iron uptake A	?	N	N
PsaA	Pneumococcal surface antigen A	++	N	N
Ply	Pneumolysin	N	?	?

Biomarker inclusion list

S. pneumoniae

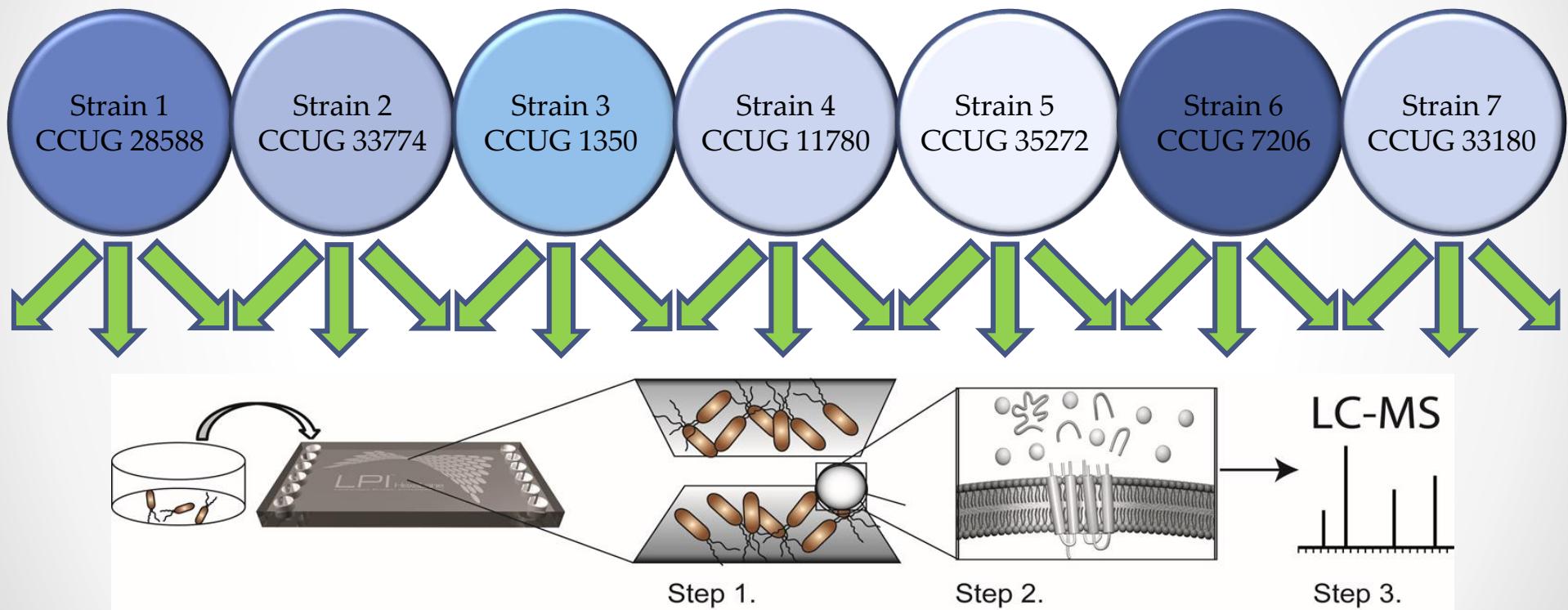
Workflow

Selection of potential biomarker peptides

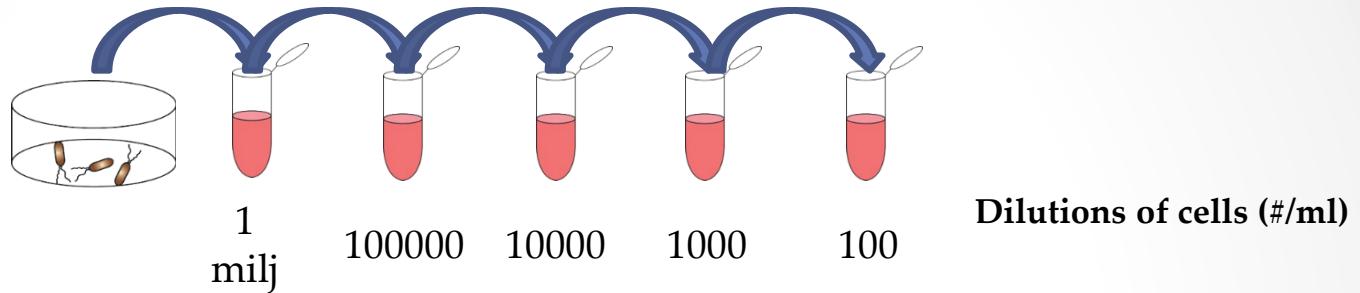


Biomarker inclusion list for *S. pneumoniae*

7 strains – including Type strain



Biomarker inclusion list for *S. pneumoniae*



Sequence	Inclusion list				Normal non-targeted			
	100	1000	100000	1 milj	100	1000	100000	1 milj
VSDVAESTGEFTSEQFEK	High	High	High	High			High	High
DAELLFAGIVGDTGR	Low	High						
YTLLAGETPAVAAAIR	High	High	High				High	
IDTFGTGTVAESQLEK	High	High	High				High	High
NGNYETAEGSEETSSEVK	High	High	High				High	High
AHIYFINSEEPSQLNDLQAFR	High	High	High				High	
ANYGVSADK		High	High				High	
AVSAILGELNGNEGK		High	High				High	High
LMDVAQPELAIVFGR		High	High					

Analysis of positive clinical samples – respiratory tract samples

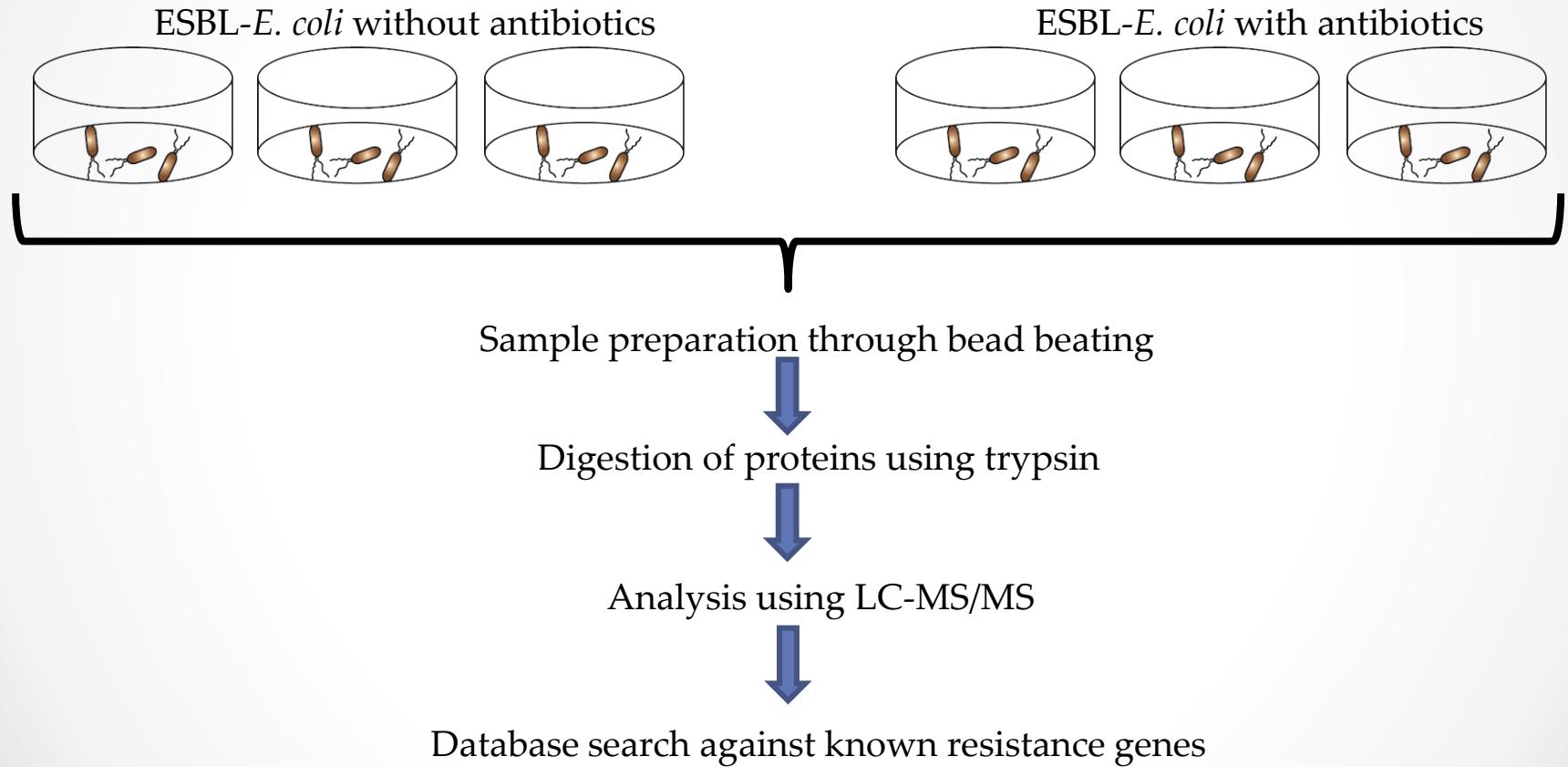
Sputum from CVID patients

	EU sample	BACT LAB IDENTIFICATION	TCUP - open	Targeted	
Clinical samples	1257	Negative	S. mitis	S. pneumoniae M. catarrhalis	
	1258	Negative	-	-	CORRECT
	1259	Negative	-	M. catarrhalis??	
	1260	Negative	-	-	CORRECT
	1261	Negative	S. mitis	S. pneumoniae M. catarrhalis	
	1262	Negative	-	-	CORRECT
	1263	H. influenzae - isolate analyzed as 1266	-	-	
	1264	H. influenzae - isolate analyzed as 1268 S. pneumoniae - isolate analyzed as 1267	H. influenzae S. pneumoniae	H. influenzae S. pneumoniae	CORRECT
	1264	H. influenzae - isolate analyzed as 1268 S. pneumoniae - isolate analyzed as 1267	H. influenzae S. pneumoniae	H. influenzae S. pneumoniae	CORRECT
	1265	Negative	-	M. catarrhalis	
Isolates from clinical samples	1266	H. influenzae	-	H. influenzae	CORRECT
	1267	S. pneumoniae	S. pneumoniae	S. pneumoniae	CORRECT
	1268	H. influenzae	H. influenzae	H. influenzae	CORRECT

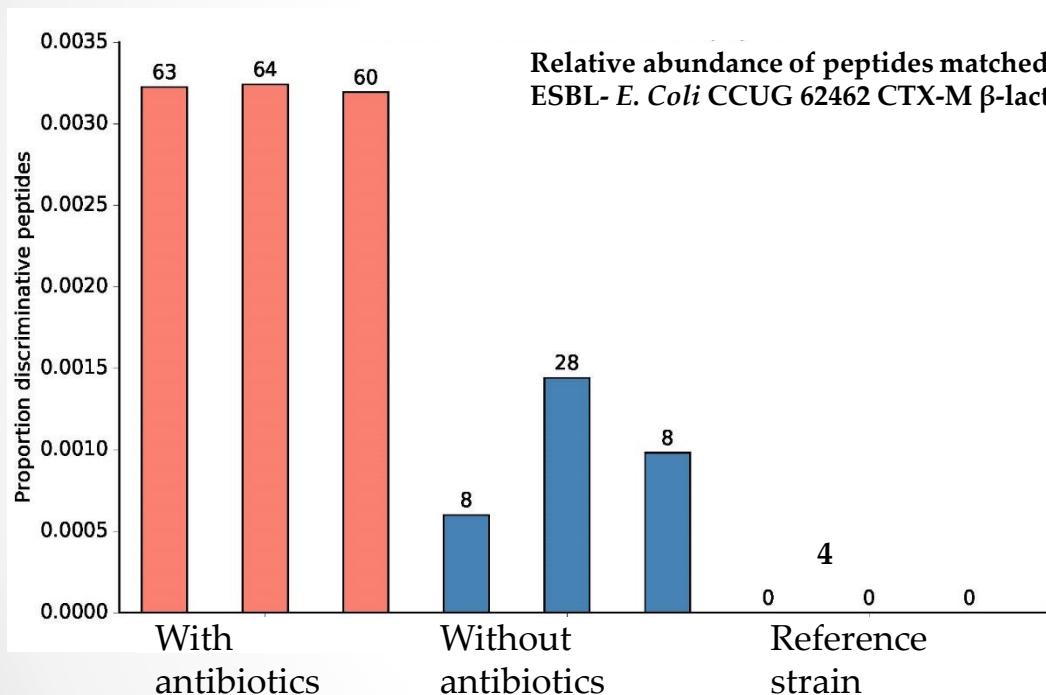
Proteotyping - Detection of anti-microbial resistance (AMR)

- WGS of urinary tract infectious bacteria
 - Type strains
 - Model strains
 - Multiresistant strains (CCUG, Sahlgrenska, Rotger)
- Inclusions lists
- Pan-genome analyses

Detection of anti-microbial resistance (AMR)



Detection of anti-microbial resistance (AMR)



- The number of peptides detected that match to CTX-M β -lactamase are significantly higher with exposure to antibiotic than without.
- These data were compared with those from analyses of a non-ESBL-*E. coli* Reference strain.
- All analyses were done in triplicate.

Experiments:
Resistance factors detected
In WGS

Conclusions

- Proteotyping provides high-resolution characterisation of bacteria;
- Proteotyping is able to detect virulence factors;
- Proteotyping is able to detect AMR factors;
- Proteotyping is applicable to clinical samples, without cultivation;
- Proteotyping requires comprehensive and accurate wgs database;
- Proteotyping requires a stable taxonomy;
- Bacteria do not cooperate with microbial systematists;
- There is more interesting work to be done



Acknowledgements

Elisabeth Inganäs, BMA
Maria Ohlén, BMA
Susanne Jensie Markopoulos, BMA
Sofia Cardew, BMA
Kent Molin, Chemist

Roger Karlsson
Francisco Salvà-Serra
Hedvig Jakobsson
Daniel Jáen-Luchoro
Lucia Gonzales-Siles
Nahid Karami
Shora Yazdanshenas
Bea Piñeiro-Iglesias
Eri (k Kristiansson
Fredrik Boulund
Elisabeth Carlsohn

CCUG Lab
CCUG Lab
CCUG Lab
CCUG Lab
CCUG Lab

CCUG - Proteomics
CCUG - Genomics
CCUG - Genomics
CCUG - Genomics
CCUG - Microbiology
SU - Microbiology / AMR
SU – Microbiology (RA)
SU - Microbiology (RA)
Chalmers Univ – Bioinformatics
Chalmers Univ – Bioinformatics
GU Proteomics Core Facility

EU – TAILORED-Treatment Project (www.tailored-treatment.eu)



ALF-Medel / Regionala FoU

