Comparative analysis of different MLST schemes for C.trachomatis with wgMLST

26th ECCMID Amsterdam, Netherlands 9 – 12 April 2016
The congress of ESCMID

Dmitriy Babenko[1], Zhanna Amirbekova[1], Anar Turmuhambetova[1], Ilya Azizov[1]
[1]Karaganda State Medical University,Karaganda, Kazakhstan.
streampost@gmail.com

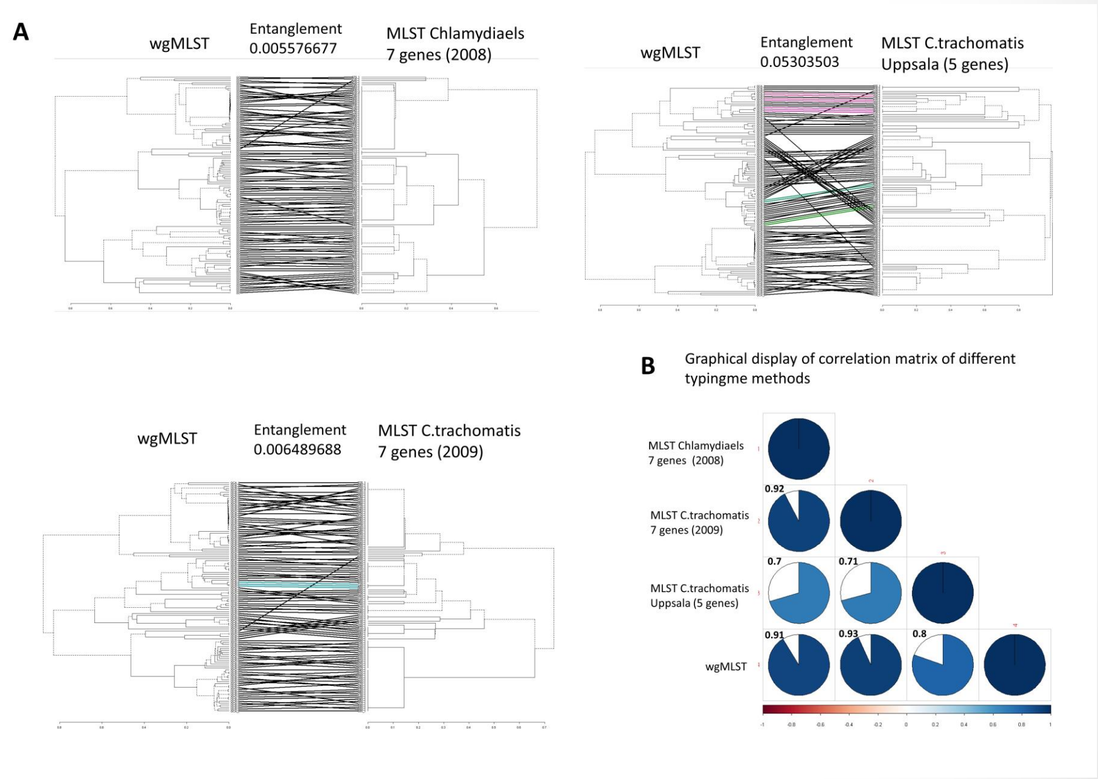
Background:
Different sequence-based typing methods have been developed to better understand the epidemiology of chlamydial diseases. Presently, three MLST schemes for C.trachomatis typing have been published, where five or seven targets with relatively high variability were used (Klint et al., 2007, Pannekoek et al., 2008, Dean et al., 2009). With development of high throughput sequencing approach, MLST have been extended to whole genome multilocus sequence typing (wgMLST) with query of hundreds/thousands of genes/loci from across the genome. The aim of the study was to compare sequence-based typing methods with wgMLST.

Material/methods:
Ad hoc wgMLST scheme with 874 loci were created based on annotated reference genome of C.trachomatis (SeqSphere software, Ridom). Genome with more 10% failed targets were removed and as a result, 118 out 128 isolates were used in analysis. Genome sequences were used to determine sequence types using three MLST schemes (http://pubmlst.org/perl/bigsdb/bigsdb.pl?db=pubmlst_chlamydiales_seqdef) with Microbial in silico typer (MIST) (Kruczkiewicz et al., 2011). Isolates with different only one locus belong to same clonal complexes. Discriminatory power and concordance between typing approaches were calculated using Simpson and adjusted Rand coefficients (Ridom EpiCompare). Tree comparing were performed with dendextend packages (Rstatistics).

Results:
Results of comparison three MLST scheme with wgMLST are presented in Table1.

Typing method	Discriminatory index (95%CI)	Mean relative difference in branch heights (comparing with wgMLST)	Correlation matrix comparing with wgMLST	Entanglement	MST statistics of clonal complexes	Concordance with wgMLST (10% cutoff – 87 alleli out 874) on clonal complex level	Concordance with wgMLST (20% cutoff – 175 alleli out 874) on clonal complex level	Concordance with wgMLST (30% cutoff – 265 alleli out 874) on clonal complex level
MLST Chlamydiales 7 genes	0.87 [0.832-0.908]	0.5732477	0.91	0.005576677	3 CC (111 samples)	adj.Rand=376 n=92	adj.Rand=0.516 n=105	adj.Rand=0.852 n =107
MLST C.trachomatis 7 genes	0.862 [0.825-0.9]	0.5278899	0.93	0.006489688	3 CC (109 samples)	adj.Rand=0.273 n=88	adj.Rand=0.353 n=101	adj.Rand=0.607 n=105
Uppsala MLST C.trachomatis 5 genes	0.953 [0.933-0.974]	0.8233071	0.8	0.05303503	6 CC (75 samples)	adj.Rand=0.846 n=62	adj.Rand=0.748 n=70	adj.Rand=0.518 n=73
wgMLST (874 genes)	0.999 [0.998-1.0]	0	1	0	10 % cutoff 14 CC (95 samples) 20 % cutoff 11CC (109 samples) 30 % cutoff 7 CC (113 samples)	-	-	-

Tree comparing and entanglement between two dendrograms based on different MLST and wgMLST and as well as correlation matrix between typing methods are shown in figure 1.



Conclusions:
The advantages of Uppsala MLST scheme with 5 targets were higher resolution power (95%) and concordance (~75-85%) with wgMLST at clonal complexes level (cutoff 10-20%) but only 59% (62-70 of 118) sample used in complexes comparing. Other MLST schemes with 7 loci demonstrated ~ 87% discriminatory ability, good concordance 85% for Chlamydiales scheme (2008) and 60% for MLST with 7 genes (2009) with wgMLST (30% cutoff), and as well as good agreement of tree topology what proves by entanglement and correlation matrix.