GERMAN FEDERAL INSTITUTE FOR RISK ASSESSMENT

BfR
Bundesinstitut für Risikobewertung

NGS analysis of rare *Campylobacter* isolates with ambiguous species differentiation

Marie-Theres Knüver, Julia Golz, Kerstin Stingl

NRL for *Campylobacter* of Germany

Species differenciation by real-time PCR

→ What do we find?

29 isolates with ambigous *mapA*/*ceuE* detection

out of 2059 *C. jejuni* and 775 *C. coli*

(Mayr et al. 2010 = Best et al. 2003 extended for *C. lari* detection; L06.32-1:2013, §64 German Feed and Food Law (LFGB))

→ around 1 % ambigous results

→ Van Rensburg et al. 2016 J. Clin. Microbiol.:

out of 1713 isolates 6 *C. coli* (0.3 % of all isolates) were ambigous in *mapA*/*ceuE* species differenciation by real-time PCR

Species differenciation by real-time PCR

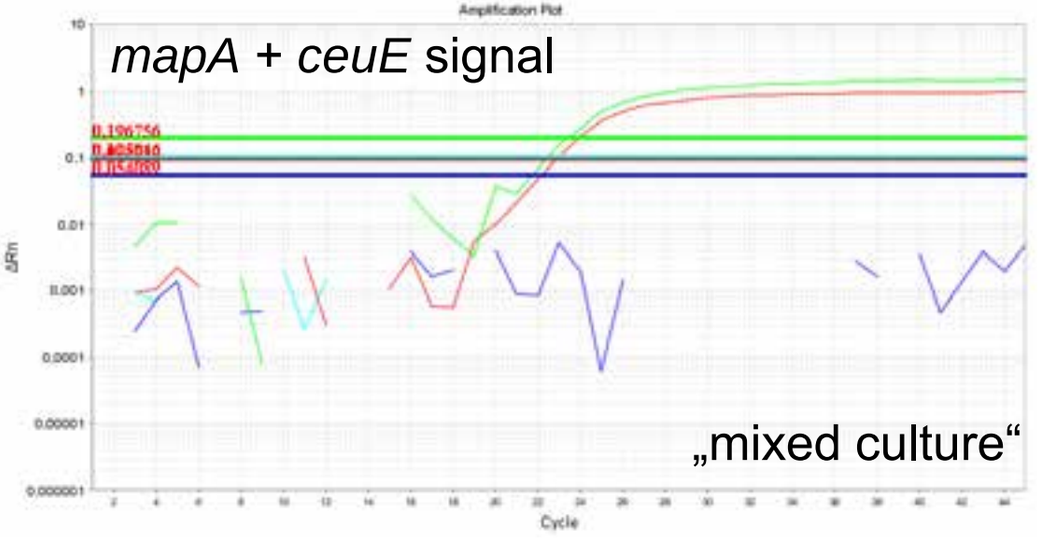
→ Where do these isolates come from?

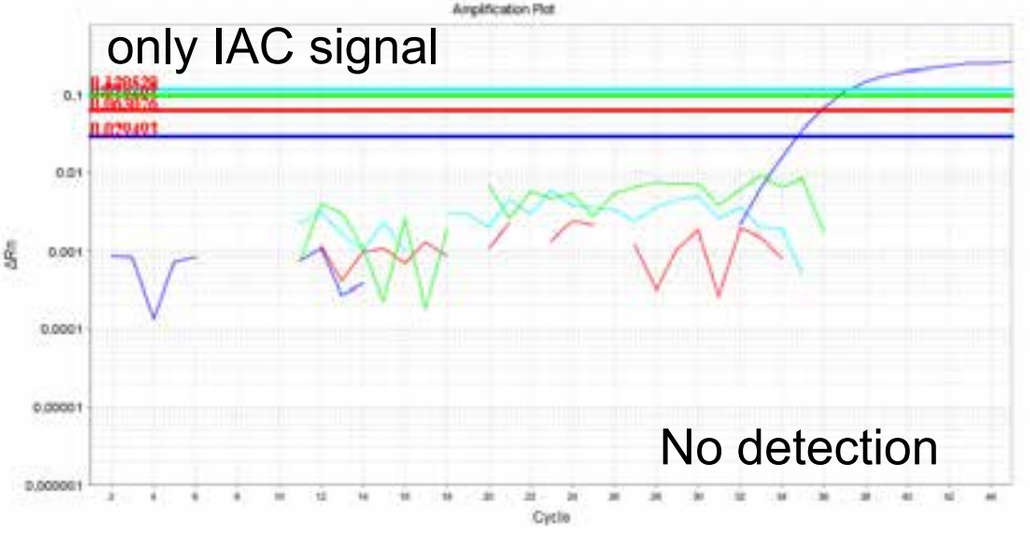
Broiler meat, eggs, turkey meat/skin/cecum, duck meat

→ What is the real-time PCR phenomeon?

„mixed“ culture = both positive for *mapA* and *ceuE* (n=27)

no Cj, Cc, Cl = negative for both *mapA* and *ceuE* (n=2)





MiSeq and PacBio sequencing

	MiSeq (BfR)	PacBio (GATC)
principle	Massively parallel sequencing (MPS)	Single molecule real-time sequencing (SMRT)
Read length	2x 300 bp	~10-20 kb
DNA required	20-100 ng	~5 µg
Instrument time	2 days	2 hours
price	80 Euro	2000 Euro

- At BfR a platform is established, providing hardware (the sequencers), optimized protocols and knowledge for the NRLs
- currently the platform is working on a common pipeline for data analysis

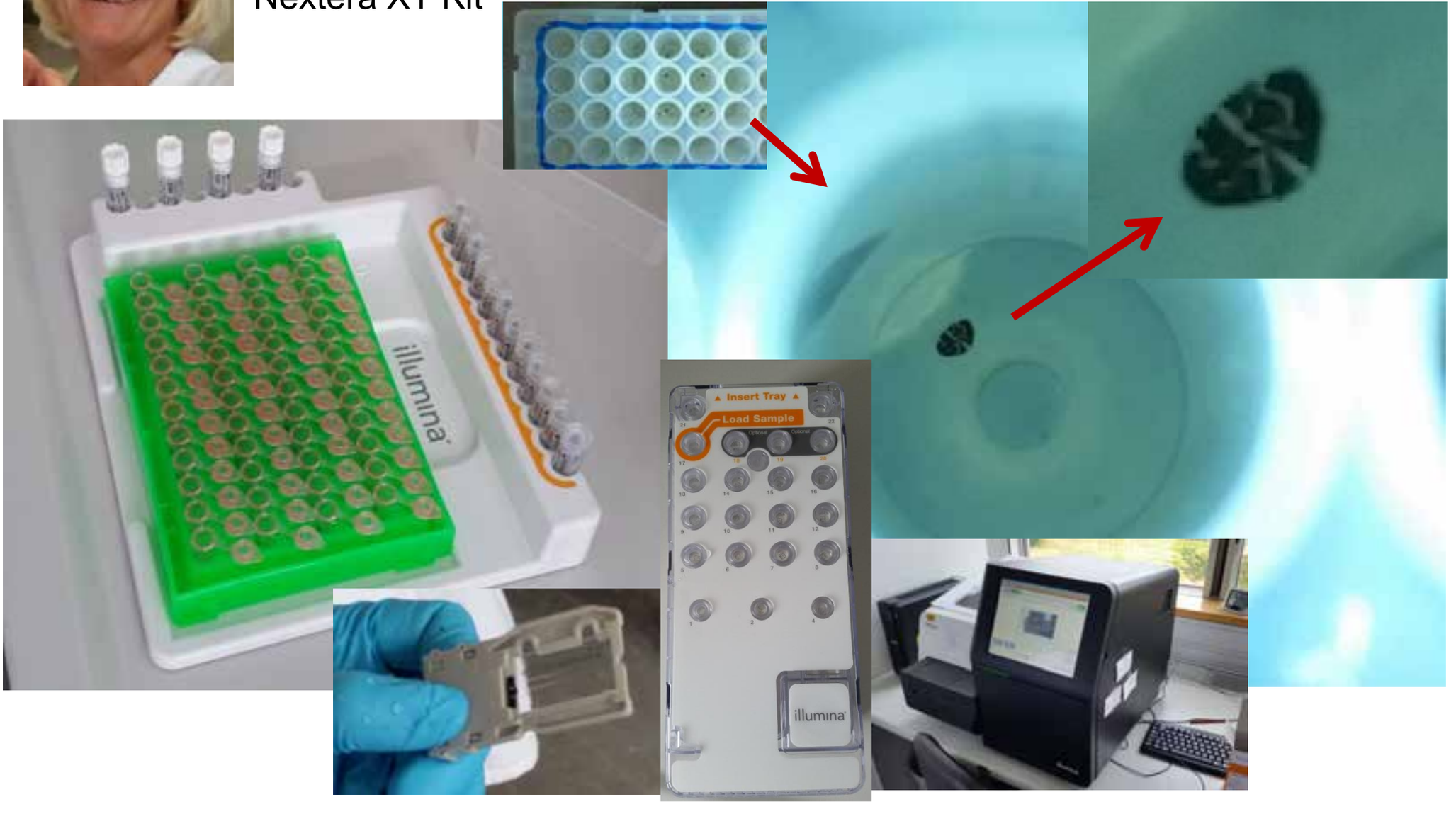
→ team of Burkhard Malorny

NGS analysis of *Campylobacter*

Illumina MiSeq of 26 of the Cj/Cc hybrids, 21 *C. coli* and 2 *C. jejuni*

Several other control strains, double sequencing for defining QC

Nextera XT Kit



NGS analysis of *Campylobacter*

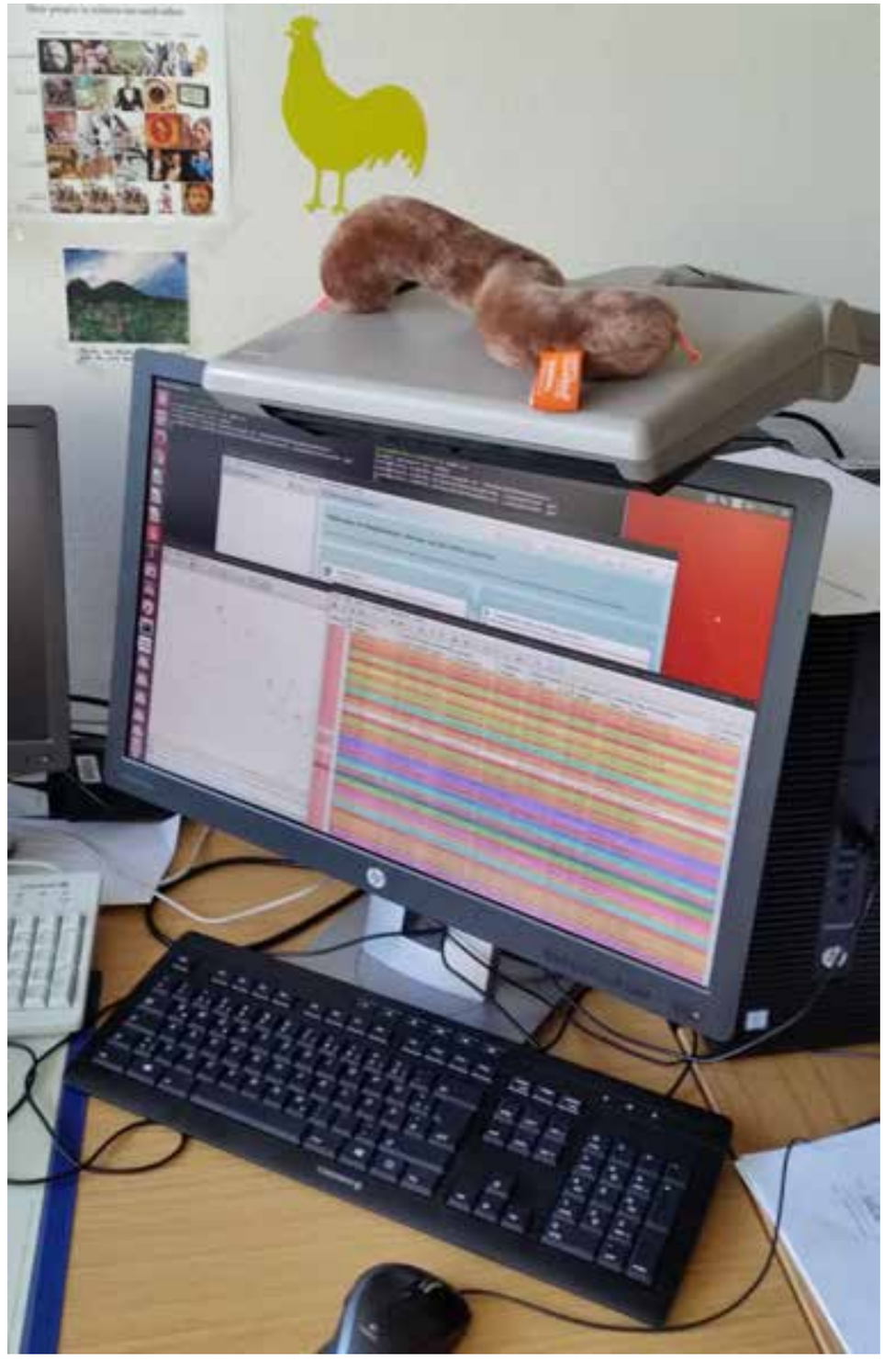
Wet lab results: fastq.gz files

Ridom Seqsphere+ 5.1.0 software
run under Ubuntu 16.04

Trimming reads to Phred score ≥ 30
Assembly via SPAdes 3.11.1

Ridom cgMLST (637 genes)
Ridom accessory genes (958 genes)

Reference Sequence (default)
is NC_002163.1.gb (NCTC 11168)



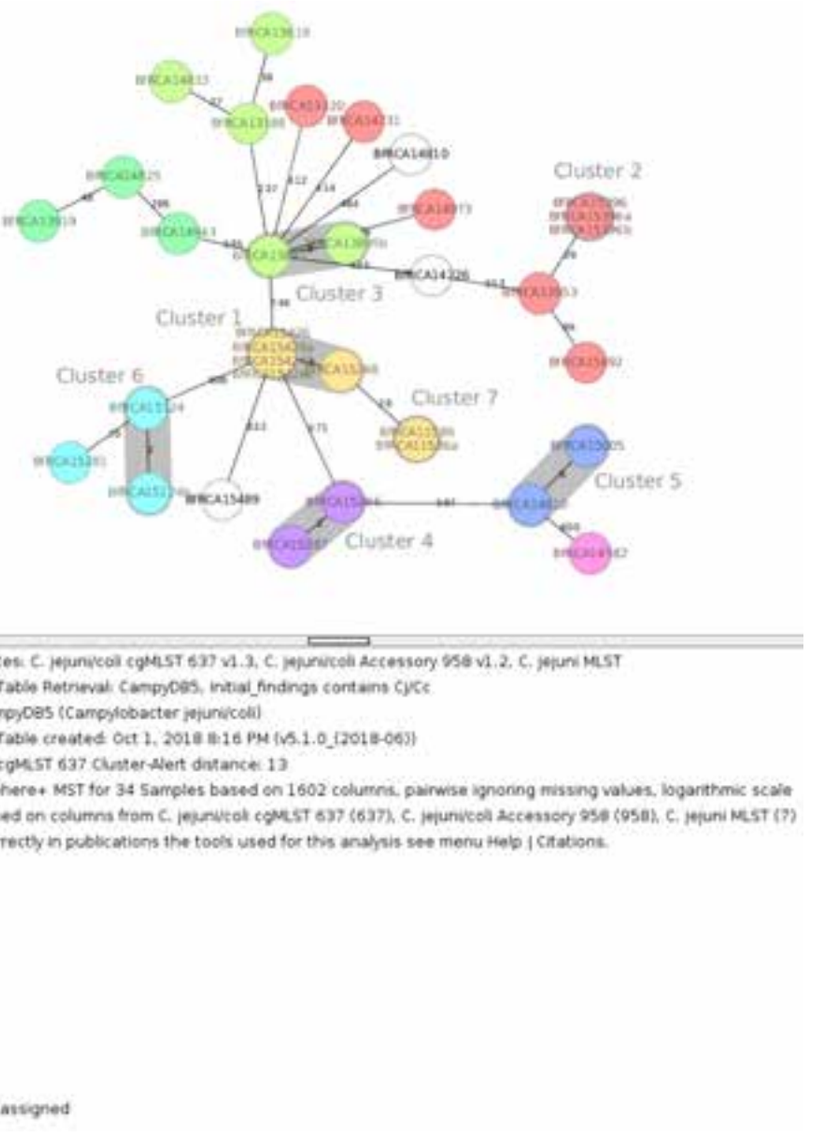
Quality control of the data

- ✓ Do reference sequences from NCBI match sequenced reference strains?
  - NCTC11168
  - 81-176
  - RM1221

  } few allele differences

- ✓ Do PacBio sequences match MiSeq sequences?
  - 7 strains were analysed by both techniques

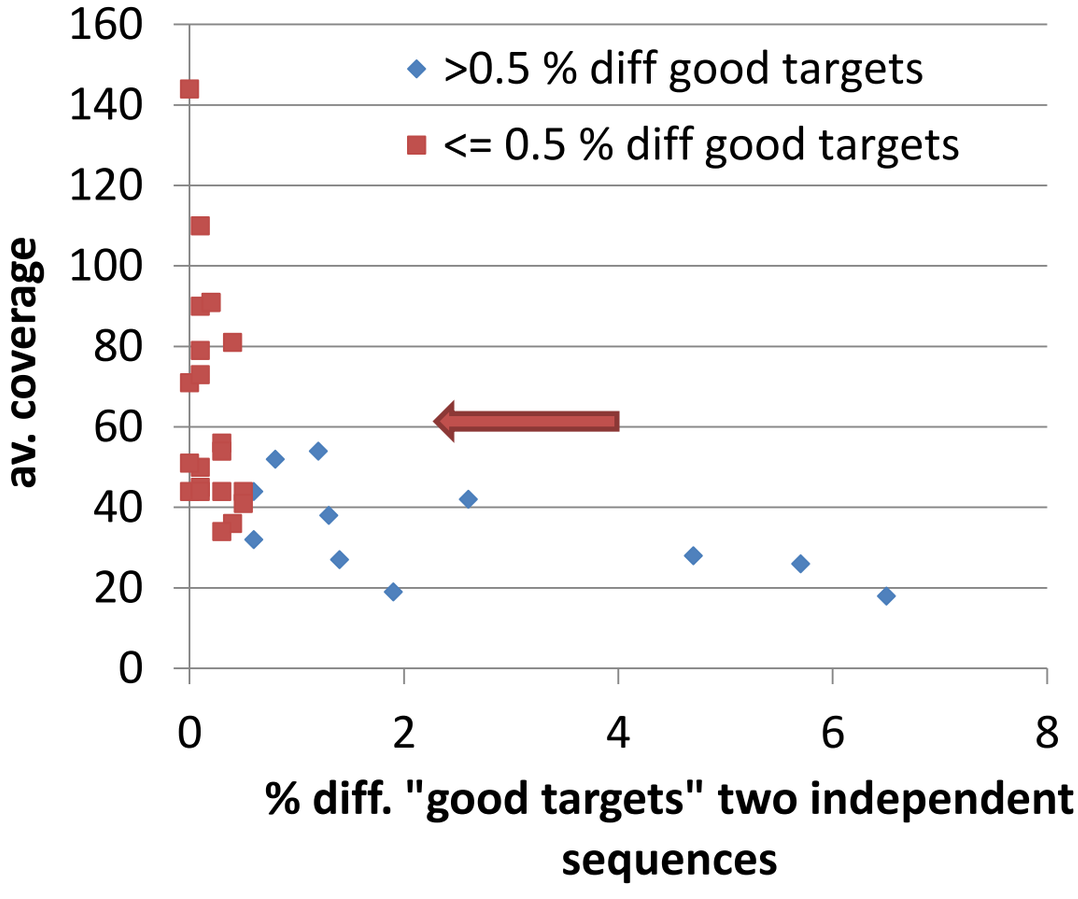
- ✓ Do they match each other if sequenced twice?
  - 28 strains were sequenced twice



Quality control of data

% good targets	diff. % good targets	Avg. Cov. (Proc., Unass.)	N50
97,8		50	178006
97,9	0,1	152	178006
97,8		131	231222
97,4	-0,4	36	84343
98,3		176	122381
97,7	-0,6	32	79334
98,5		180	154048
92	-6,5	18	16133
97,7		79	231120
97,8	0,1	180	231088

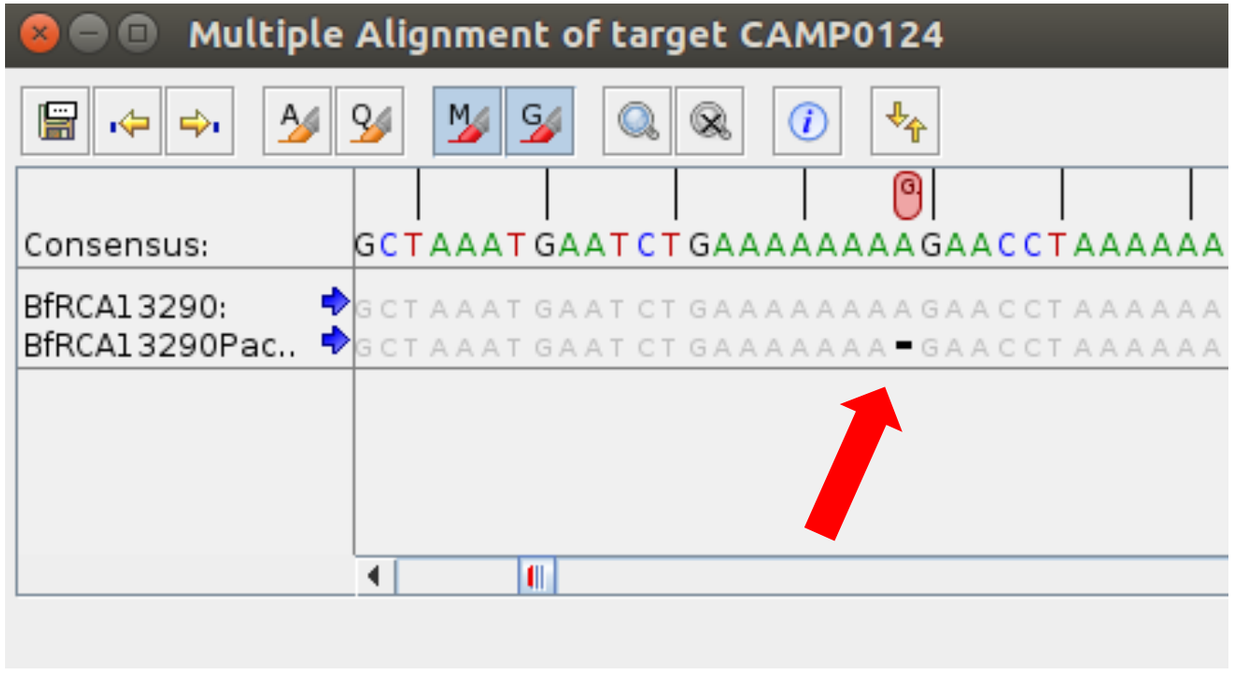
Examples of compared sequence results



→ data might indicate that an av. coverage of 60 is sufficient for cgMLST analysis by Ridom Seqsphere+ (with the additional pipeline QC parameters)

Quality control of the data

	av. Cov. MiSeq	diff. % good targets
PacBio1	180	1,7
PacBio2	180	0,8
PacBio3	121	2,1
PacBio4	99	6,3
PacBio5	180	1
PacBio6	180	3,4
PacBio7	144	0,7

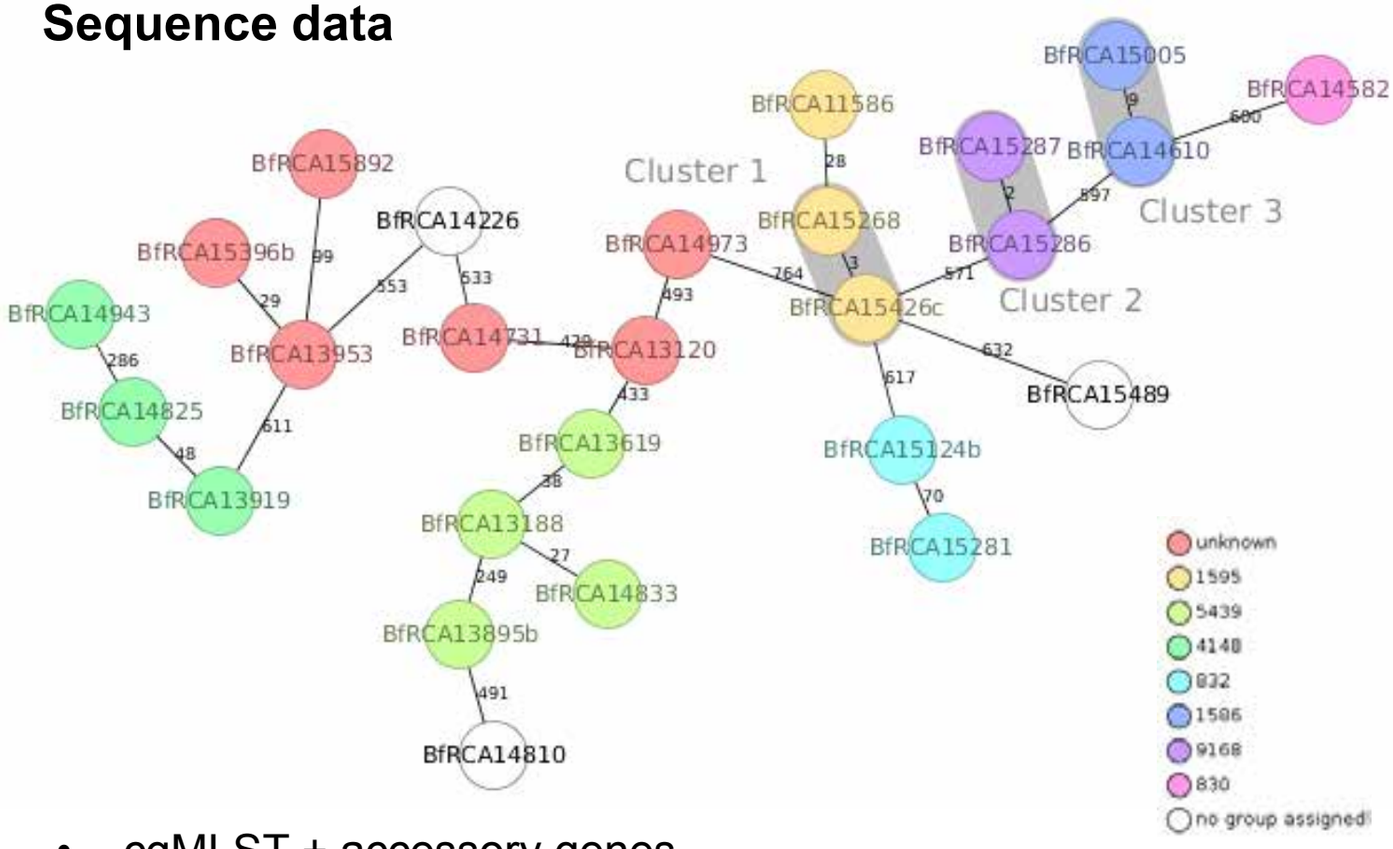


- although PacBio is superior in terms of gene order and assembly of 1 whole chromosome plus epichromosomal elements, insertions and deletions are frequently found

→ hybrid assembly with PacBio and MiSeq data (future task)

Are the Cj/Cc hybrids phylogenetically linked?

Sequence data

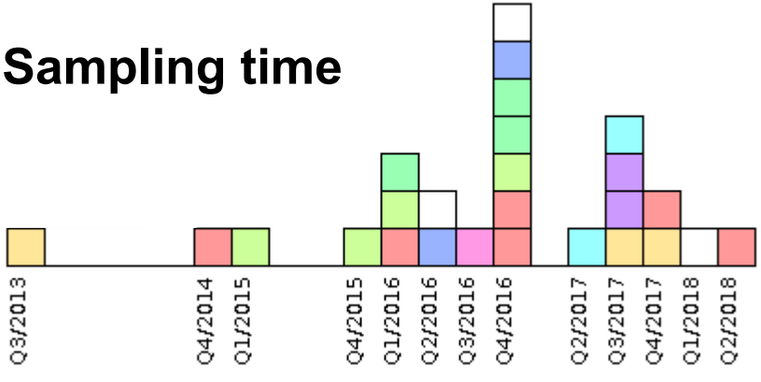


Sampling location

- cgMLST + accessory genes
- pairwise ignore missing values → more differences possible!
- 3 “clusters“ (à 2 strains); rest of the strains is phylogenetically unlinked

→ Multiple events as cause for ambigous species detection

Sampling time

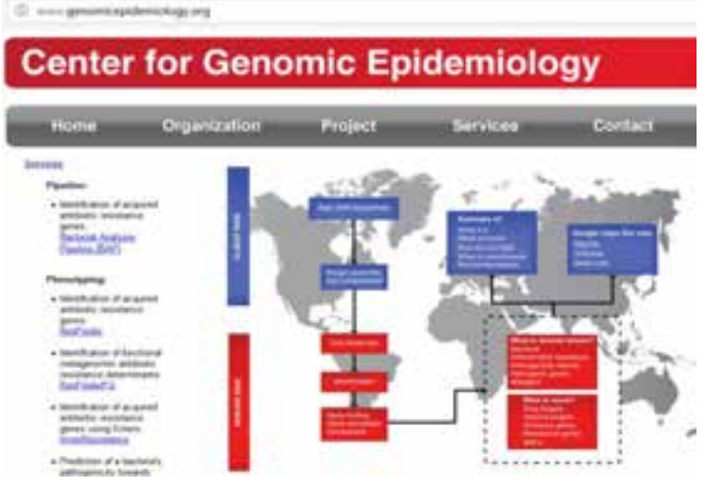


What species do the isolates with ambigous *mapA/ceuE* belong to?

Kmer Finder 3.0 by CGE genomics webpage (Kmer size 16)

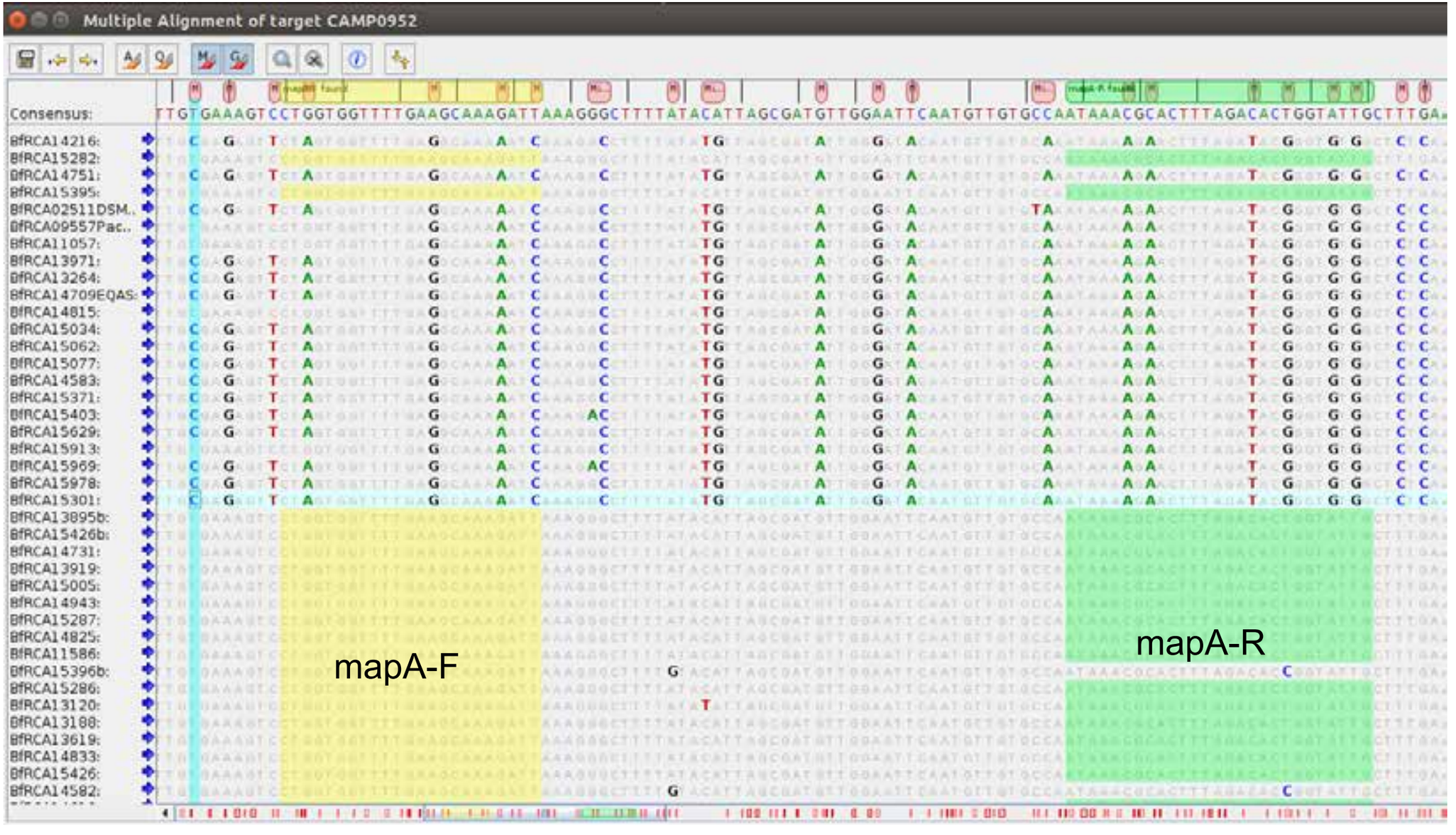
query_coverage [%]: is the percentage of input query Kmers that match the template

C. jejuni introgression	RT-PCR *C. coli*	RT-PCR Cj/Cc hybrid
0%	8	1
< 3%	10	9
3-6 %	2	1
>6 %	1	15



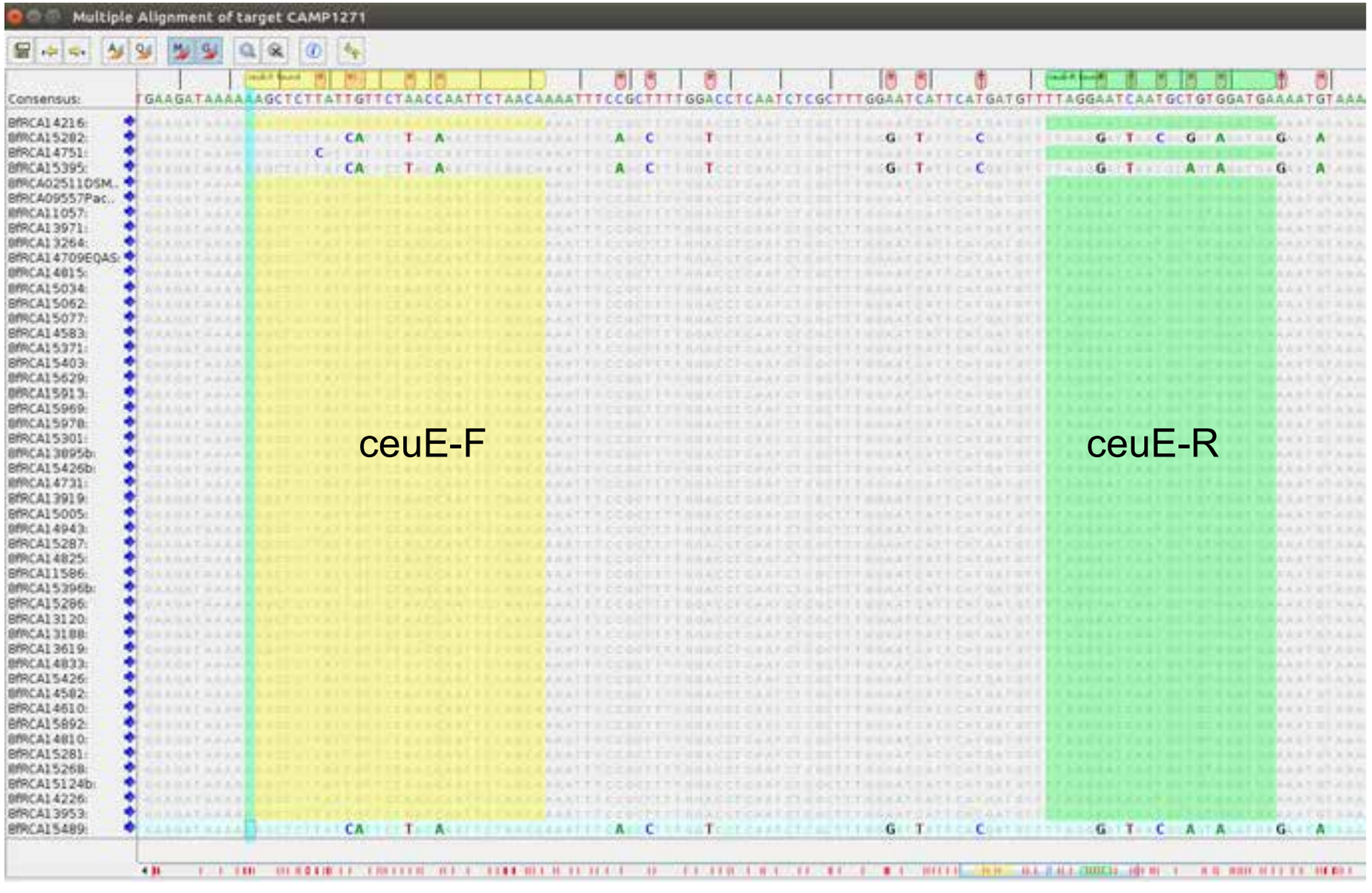
→ all isolates are *C. coli* with 0-14.4% *C. jejuni* introgression

What happened at *mapA*?



- Adjacent genes also implicated in recombination event – in some isolates Cj sequence spans over 4 genes

What happened at *ceuE*?



- 15489 has *ceuE* form *C. jejuni*; the 14973 has a *ceuE* allele below the threshold of Ridom

Could the Cj/Cc hybrids be correctly identified via other PCRs?

Gene target	function	Species	Ridom scheme	Discrimi-natory in "hybrids"	references
glyA	serine hydroxymethyl-transferase	*C. coli*	core	yes	Wang 2002; LaGier 2004
hipO	hippurate hydrolase	*C. jejuni*	accessory	yes	Wang 2002; LaGier 2004, Toplak 2012
cpn60	chaperone (*groEL*)	*C. jejuni/ C. coli*	core	partial	Chaban 2009
cadF	Fibronectin-binding protein	*C. jejuni/ C. coli*	accessory	yes	Toplak 2012; Shams 2016
lpxA	N-actylglucosamine transferase	*C. jejuni/ C. coli*	accessory	yes	Klena 2004
ccoN	cytC subunit	*C. jejuni/ C. coli*	core	partial	Toplak 2012
mapA	Outer membrane protein (MOMP)	*C. jejuni*	accessory	no	Best 2003; Mayr 2010
ceuE	enterochelin transporter, substrate-binding	*C. coli*	accessory	no	Best 2003; Mayr 2010

- Results on phylogenetic tree based on gene sequence; oligo annealing sites have to be checked
- Wang et al. 2002 multiplex-PCR identifies *C. coli* (*glyA*) in wet lab

Are other core *C. jejuni* genes, never found in *C. coli*, recombined in the Cj/Cc hybrids?

- 48 genes were defined which belong to the *C. jejuni* core and are never found in *C. coli* (Lefébure et al., 2010; Meric et al., 2014)
- we could detect introgression of *C. jejuni* in 8 of these genes
- 15 of the Cj/Cc hybrids and 2 of the *C. coli* harboured core *C. jejuni*

	core (L)	core (L)	core (L)	core (M)	core (L/M)	core (M)	core (L)	core (M)
Initial_findings	Cj0004c	Cj0005c	Cj0037c	Cj0135	Cj0327	Cj1062	Cj1589	Cj1714
C. jejuni	6	18	10	1	1	9	12	1
C. jejuni	4	23	76	1	7	36	? (failed)	5
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	8	? (not fou	? (failed)	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
C. coli	2	15	? (not fou	? (not fou	? (not foun	56	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	29	1	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	1
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	1
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	16	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	3	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	13	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	18	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	1
Cj/Cc-Hybrid	? (not fou	? (not fou	21	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	3	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	*? (not fou*	*? (not fou*	*? (not fou*	*? (not fou*	*? (not found*	*? (not fou*	*? (not fou*	*? (not found)*
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	23	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	23	? (not fou	? (not found)
Cj/Cc-Hybrid	*? (not fou*	*? (not fou*	*16*	*? (not fou*	*? (not found*	*? (not fou*	*? (not fou*	*? (not found)*
Cj/Cc-Hybrid	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid	? (not fou	? (not fou	? (failed)	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid*	? (not fou	? (not fou	? (not fou	? (not fou	? (not foun	? (not fou	? (not fou	? (not found)
Cj/Cc-Hybrid*	? (not fou	? (not fou	22	? (not fou	? (not foun	13	? (not fou	? (not found)

Conclusions and perspective

- *mapA*/*ceuE* species differentiation is ambigous for ~1 % of Cj/Cc isolates but does not provide false identification

- NGS analysis revealed them to be *C. coli,* which was cofirmed by Wang et al. 2002 (*glyA*)

→ Good decision to add both PCR options as Annex to ISO 10272

- QC analysis suggested that an average coverage of ≥60 might be suitable (Ridom pipeline)
- PacBio data should be „cured“ for insertion/deletions by MiSeq data (future task to use hybrid assembly such as Unicycler)

- Further characterization of the *C. coli* with hudge amount of *C. jejuni* introgression
  - Are they special or just found by chance?
  - Do they have extended DNA uptake capacity?
  - Is there a trigger for DNA exchange by natural transformation?

GERMAN FEDERAL INSTITUTE FOR RISK ASSESSMENT

Thanks to…
the Federal Laender laboratories
for isolating *Campylobacter*

BfR NGS platform
Maria Borowiak & Burkhard Malorny

NRL team:
Christiane Buhler
Britta Kraushaar
Maja Thieck
Petra Vogt
Imke Wulsten

BfR
Bundesinstitut für Risikobewertung

Federal Ministry of Education and Research

Thank you for your attention

Kerstin Stingl

Phone +49 30 - 184 12 - 2135

kerstin.stingl@bfr.bund.de ● www.bfr.bund.de/en