

Centre National de Référence des Campylobacters  
et Hélicobacters  
CHU de Bordeaux

**Bilan de la mise en place par NGS de l'étude du résistome et de  
l'épidémiologie des infections à Campylobacters en France**

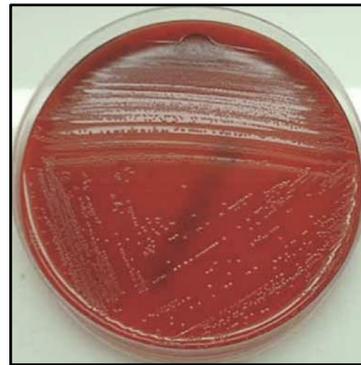
Quentin Jehanne et Philippe Lehours



# Analyses de routine sur les Campylobacters au CNRCH jusqu'à fin 2023



Milieu de transport



Culture



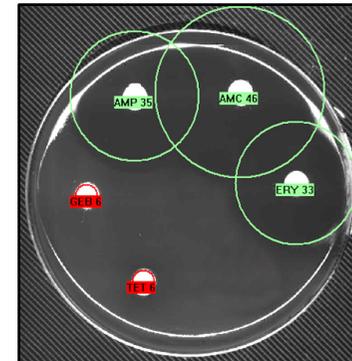
Congélation -80°C



MALDI-TOF MS  
(Bruker Daltonics)



ATB<sup>ame</sup> (diffusion en disque, Etest)



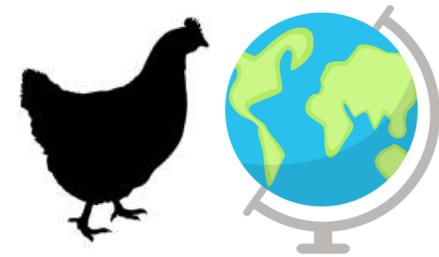
# Limites rencontrées



Espèce non identifiée  
par MALDI-TOF?



Mécanismes de  
résistance?



Provenance des  
souches (sources?)

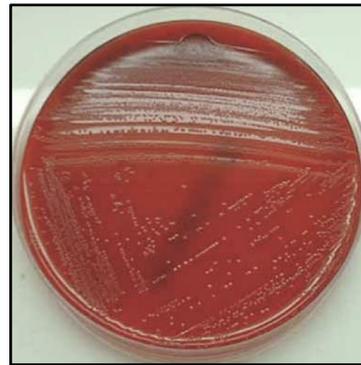
# Analyses de routine sur les Campylobacters au CNRCH depuis janvier 2024



Milieu de transport

Pas d'ATBme  
réalisé

Fiche CNR



Culture

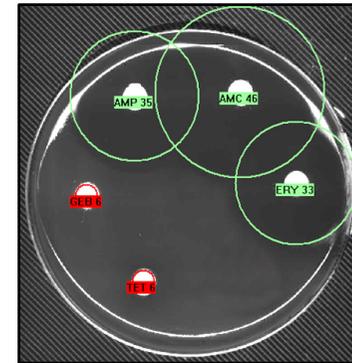
Congélation -80°C



MALDI-TOF MS  
(Bruker Daltonics)



ATB<sup>ame</sup> (diffusion en disque, Etest)



# Analyses de routine sur les Campylobacters au CNRCH depuis janvier 2024



Milieu de transport

ATBme  
réalisé

Fiche CNR



Culture

Congélation -80°C



MALDI-TOF MS  
(Bruker Daltonics)

NGS (Novaseq 6000)

- plateforme NGS au CHU Henri Mondor (APHP, Créteil)
- deux envois de suspensions/semaine



## En pratique

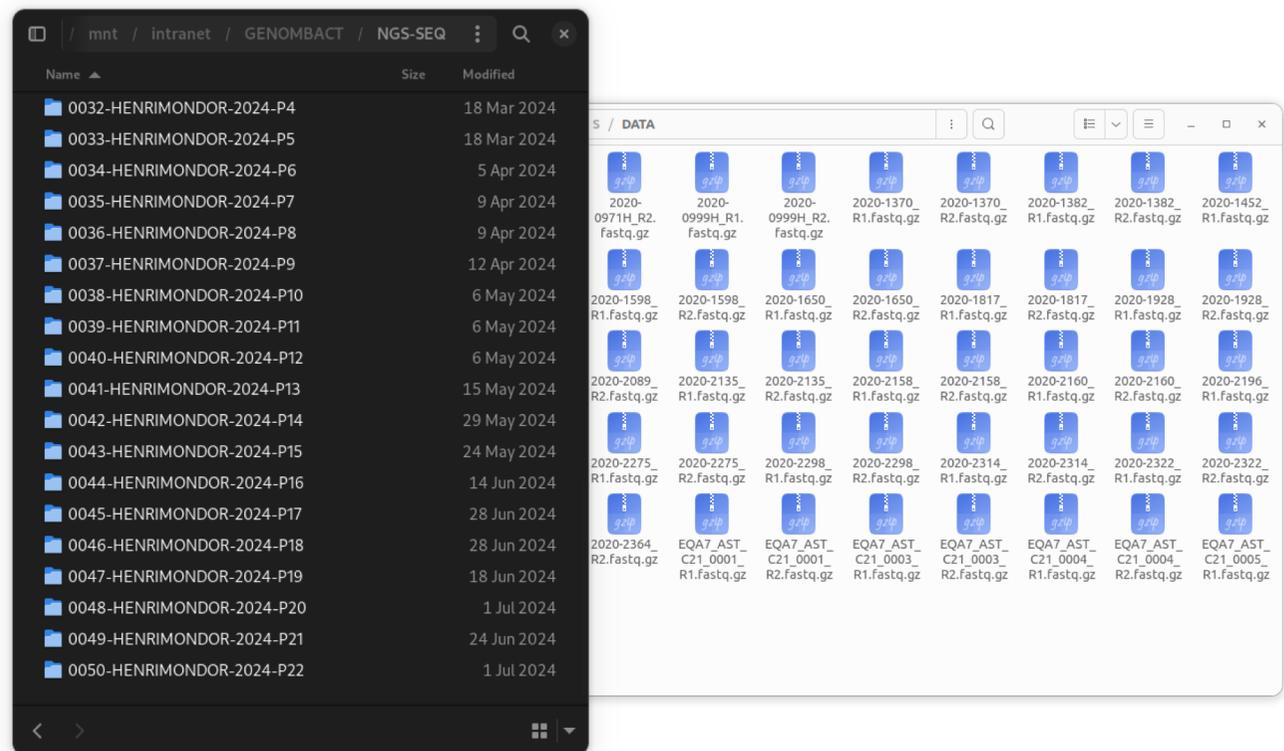
- Sur budget du CNR
  - › non facturé
  
- Délai de rendu de résultat
  - › 18 jours environ
  - › + 7 jours si nécessité de confirmer des discordances phénotype/résistome
  
- EEQ depuis organisé par l'ECDC pour les pipelines bioinformatiques
  
- Résultat validé disponible sous format mail sécurisé (mmsanté, apycrypt)



## Les analyses bio-informatiques en détails

# Données du CHU Henri Mondor reçues et traitées automatiquement

- Envoi des données brutes de séquençage par un système de transfert sécurisé (SFTP) entre l'APHP et le CHU de Bordeaux.



# Pipeline d'analyses bio-informatiques

- Développée avec Python et Bash.
- Utilisable sous le système d'exploitation Linux.
- Se lance toutes les nuits si des séquences sont disponibles.

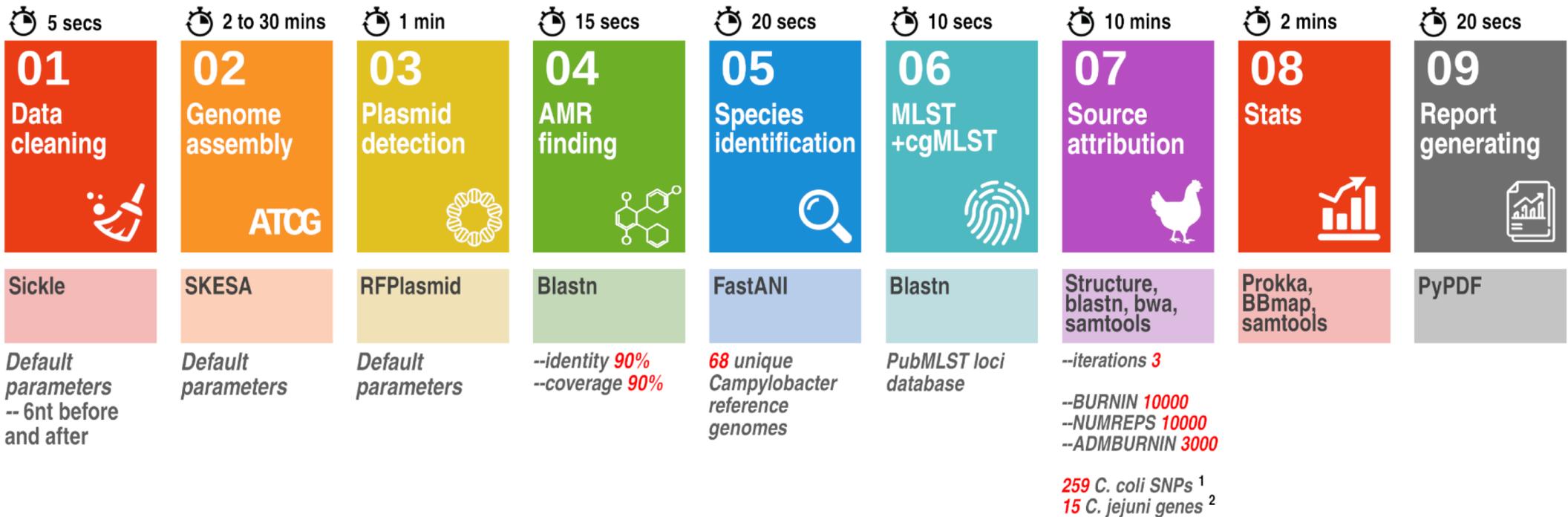
```
1 #!/bin/bash
2
3 basedirectory=${dirname "$0"}
4 source $basedirectory/"ini-file-parser.sh
5 process_ini_file $basedirectory"/parameters.conf"
6
7 VER_BG="\033[43m"
8 BLU_BG="\033[44m"
9 PRP_BG="\033[45m"
10 NC="\033[0m"
11
12 inputDirectory=$1
13 saveDirectory=$2
14 dataFile=$3
15
16 if [[ "$inputDirectory" == "" ]]
17 then
18     inputDirectory=$basedirectory
19 fi
20
21 if [[ "$saveDirectory" == "" ]]
22 then
23     saveDirectory=$basedirectory
24 fi
25
26 if [[ "$dataFile" == "" ]]
27 then
28     dataFile=$basedirectory"/dataFile.sh"
29 fi
30
31 outputDirectory=$basedirectory"/output"
32 rm -rf $outputDirectory
33 mkdir $outputDirectory
34
35 pythonScripts=$basedirectory"/pythonScripts"
36 py_rename=$pythonScripts"/py_rename.py"
37 py_fastaFilter=$pythonScripts"/py_fastaFilter.py"
38 py_resFinder=$pythonScripts"/py_resFinder.py"
39 py_mlst=$pythonScripts"/py_mlst.py"
40 py_cgmlst=$pythonScripts"/py_cgmlst.py"
41 py_sourceAttribution=$pythonScripts"/py_sourceAttribution.py"
42 py_statistics=$pythonScripts"/py_statistics.py"
43 py_report=$pythonScripts"/py_report.py"
44 logsPrep=$saveDirectory"/logsPrep.sh"
45
46 if [[ "$steps_ftp_download" == "yes" ]]
47 then
48     echo "date +%D"
49     ./ftp_download_files.sh >> $logsPrep 2>&1
50 fi
51
52 if [[ "$steps_rename_dl" == "yes" ]]
53 then
54     python3 $py_rename $inputDirectory $paths_information_file >> $logsPrep 2>&1
55 fi
56
57 if [ ! -d $inputDirectory ]
58 then
59     echo "Input directory $1 does not exists." >> $logsPrep 2>&1
60     exit 9999
61 fi
```

```
26 glms_folder= /mnt/intranet/GLIMSRES/
27 ariba= ariba
28 sickle= /home/qjehanne/pipeline/tools/sickle/sickle
29 spades= /home/qjehanne/pipeline/tools/SPAdes-3.15.5/spades.py
30 skesa= /home/qjehanne/pipeline/tools/SKESA/skesa
31 rfplasmid= /home/qjehanne/pipeline/tools/RFPlasmid/rfplasmid.py
32 blastn= blastn
33 blastx= blastx
```

```
qjehanne@cnrchws:~/pipeline — /bin/bash ./exec_skesa.sh
>> CNRCH CAMPYLOBACTER ANALYSIS PIPELINE <<
RUN "0028-HENRIMONDOR-2024-P0" ISOLATE 01789 (4/10)
02/09/24 14:09:26 Step 1 Data cleaning ...
02/09/24 14:10:00 Step 2 Genome assembly ...
02/09/24 14:17:31 Step 3 Plasmids detection ...
02/09/24 14:19:13 Step 4 Resistance finding ...
02/09/24 14:19:34 Step 5 Species identification ...
02/09/24 14:20:09 Step 6 Molecular typing ...
02/09/24 14:20:15 Step 7 Core genome molecular typing ...
02/09/24 14:26:03 Step 8 Source attribution ...
02/09/24 14:30:54 Step 9 Data statistics ...
02/09/24 14:42:20 Step 10 Generating results file ...
```



# Pipeline d'analyses bio-informatiques

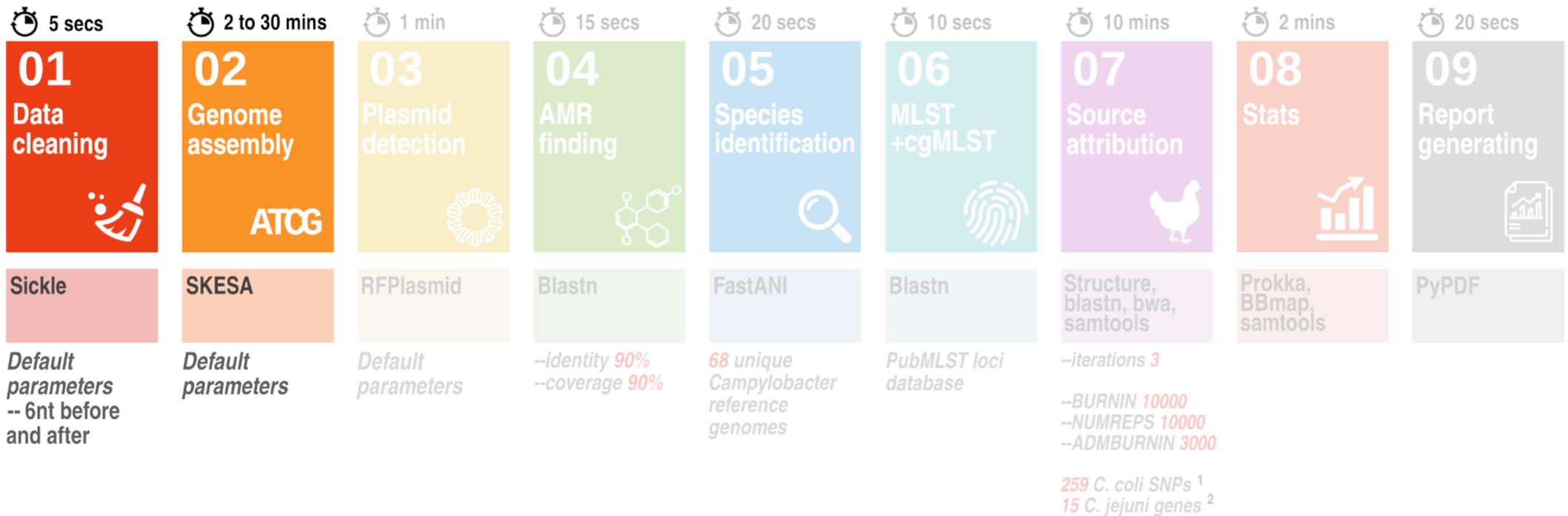


# Pipeline d'analyses bio-informatiques



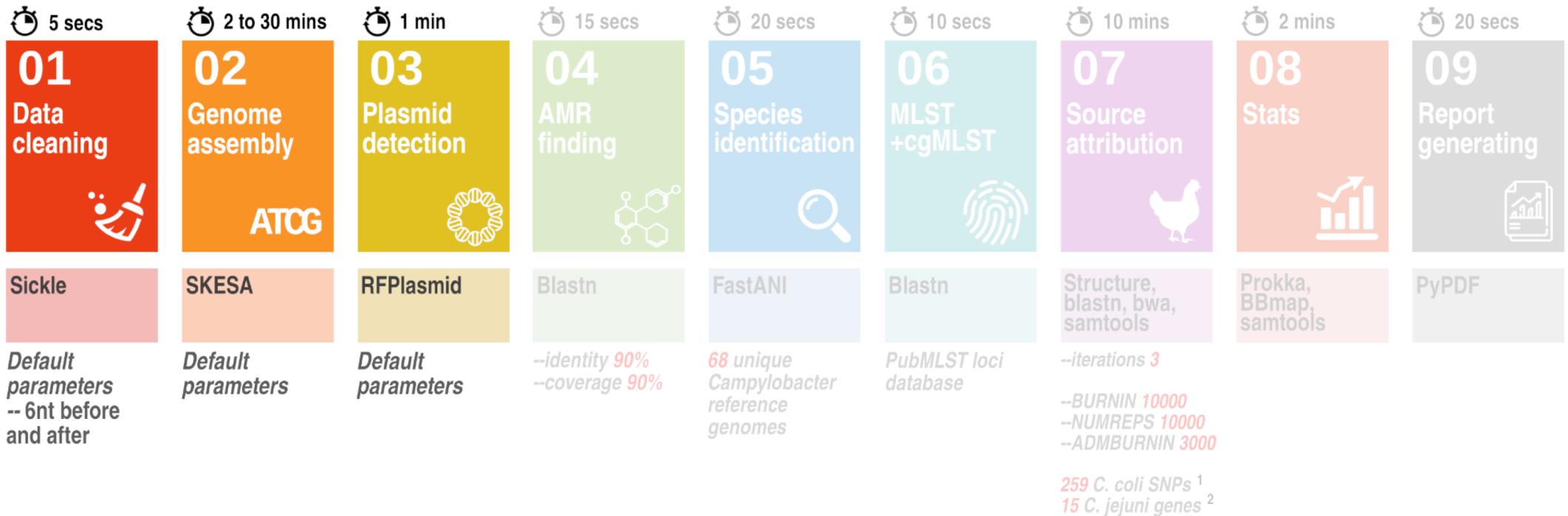
→ Suppression des séquences de mauvaise qualité.

# Pipeline d'analyses bio-informatiques



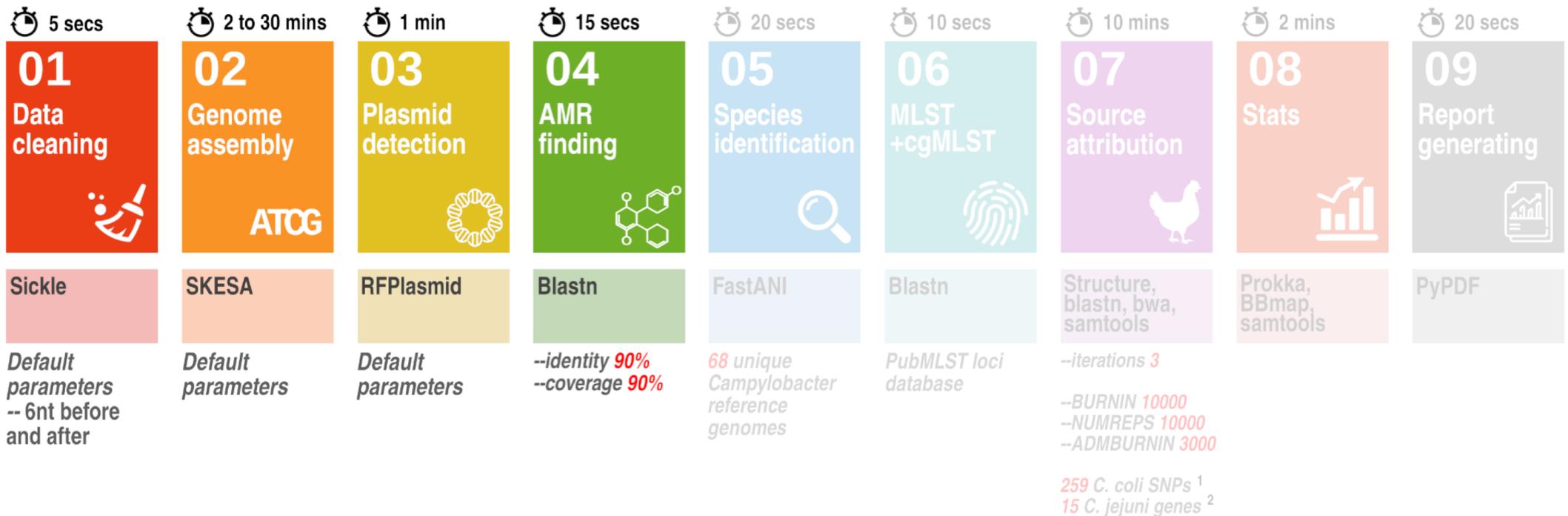
→ Génération du génome de la souche par assemblage.

# Pipeline d'analyses bio-informatiques



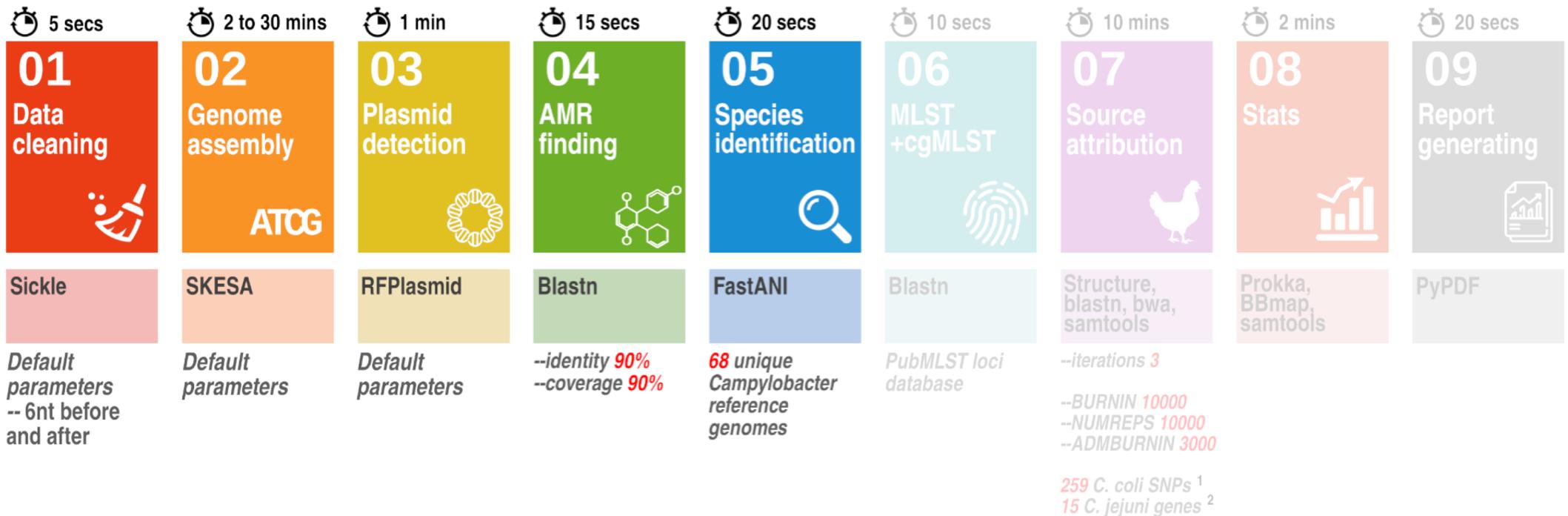
→ Détection des plasmides.

# Pipeline d'analyses bio-informatiques



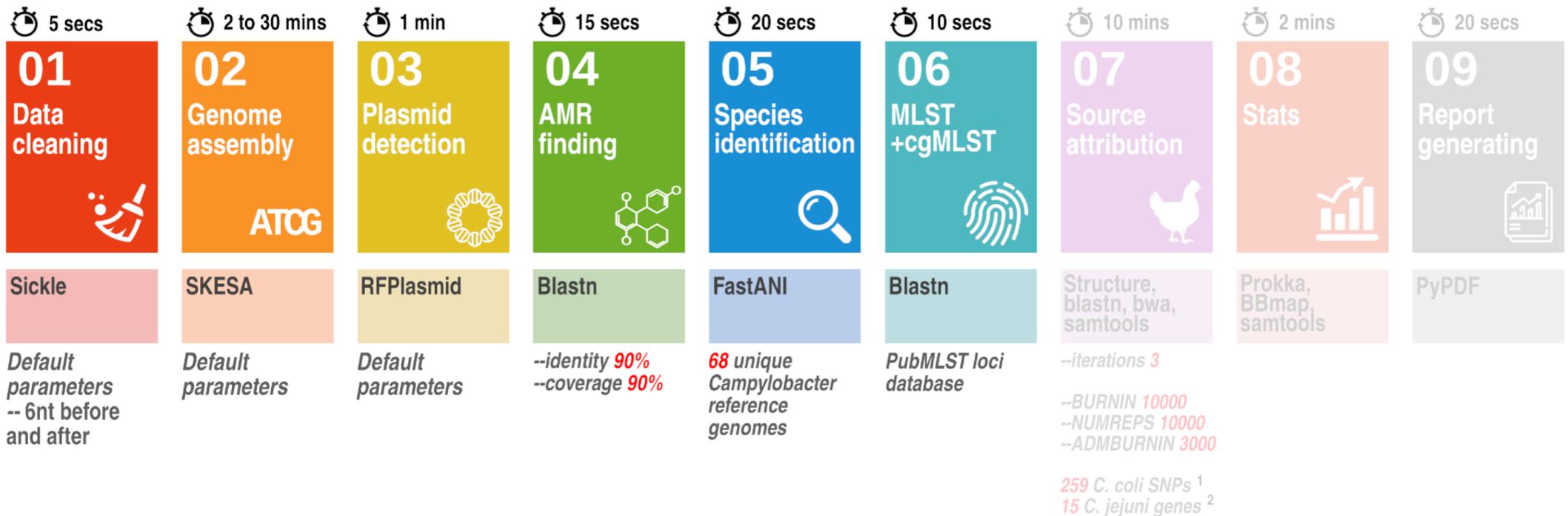
→ Détection de la résistance aux antibiotiques à partir de plusieurs bases de données de mutations et de gènes (NCBI, ResFinder, CARD, CNRCH, ...).

# Pipeline d'analyses bio-informatiques



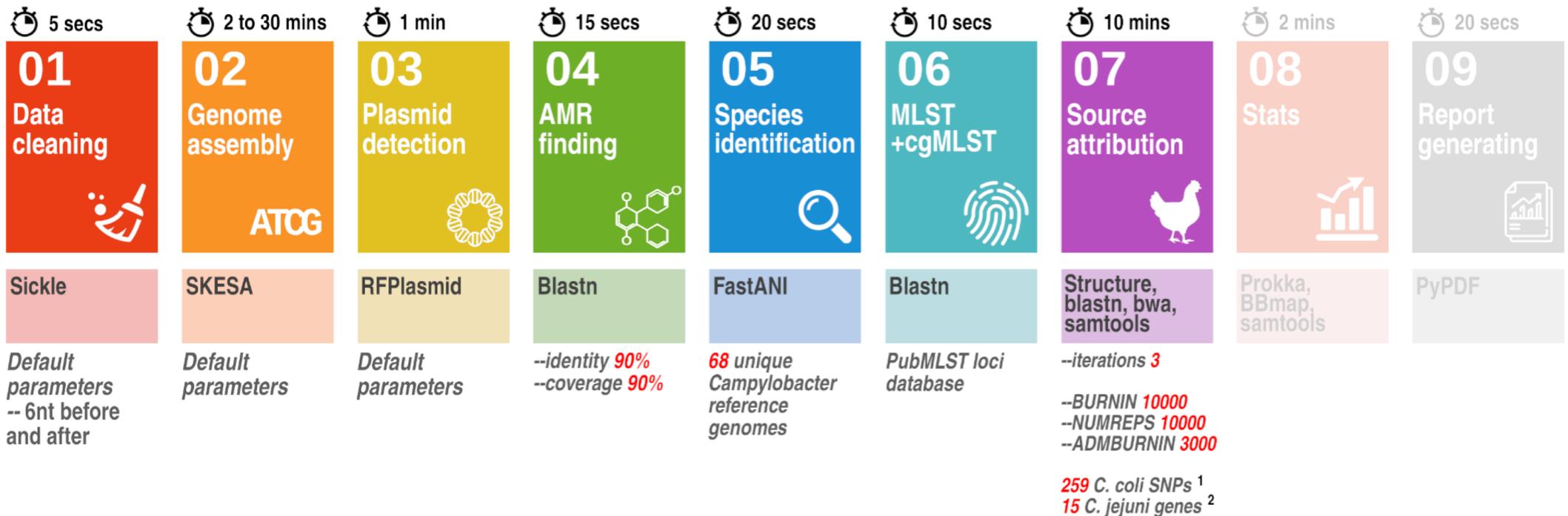
→ L'espèce est identifiée en comparant le génome à 129 références de *Campylobacter*, *Helicobacter* et *Aliarcobacter*.

# Pipeline d'analyses bio-informatiques



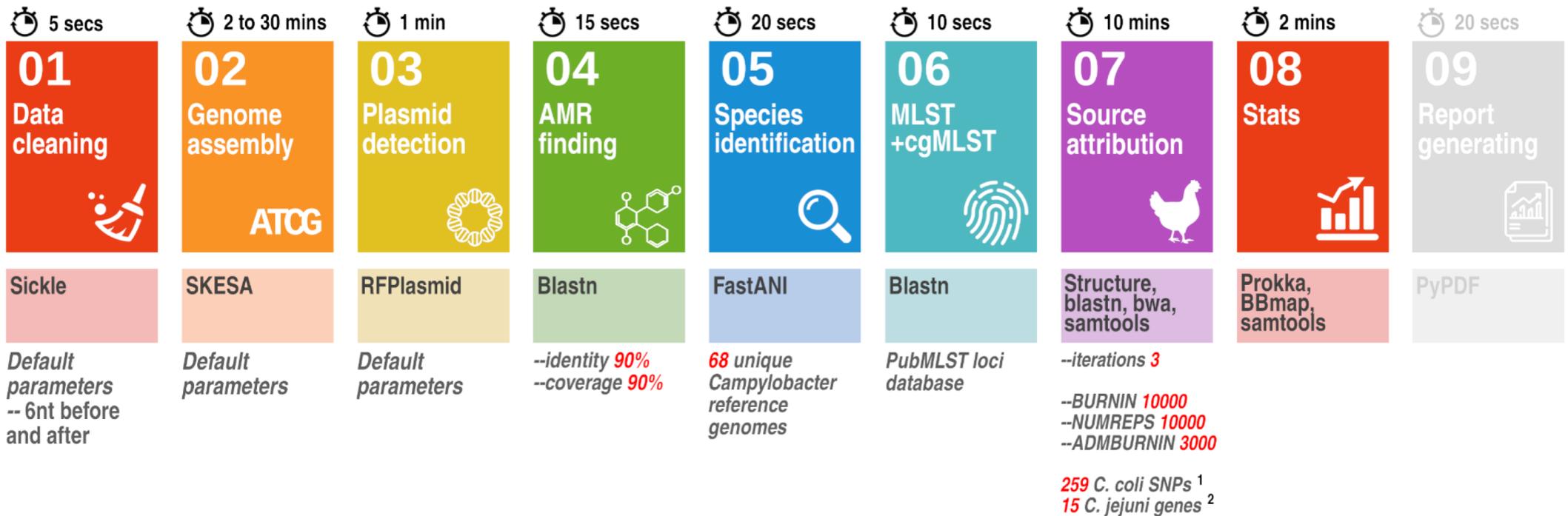
- Typage moléculaire de la souche par la méthode MLST (Dingle *et al.* 2001).
- Utilisation de la base de données PubMLST pour la MLST et le cgMLST.

# Pipeline d'analyses bio-informatiques



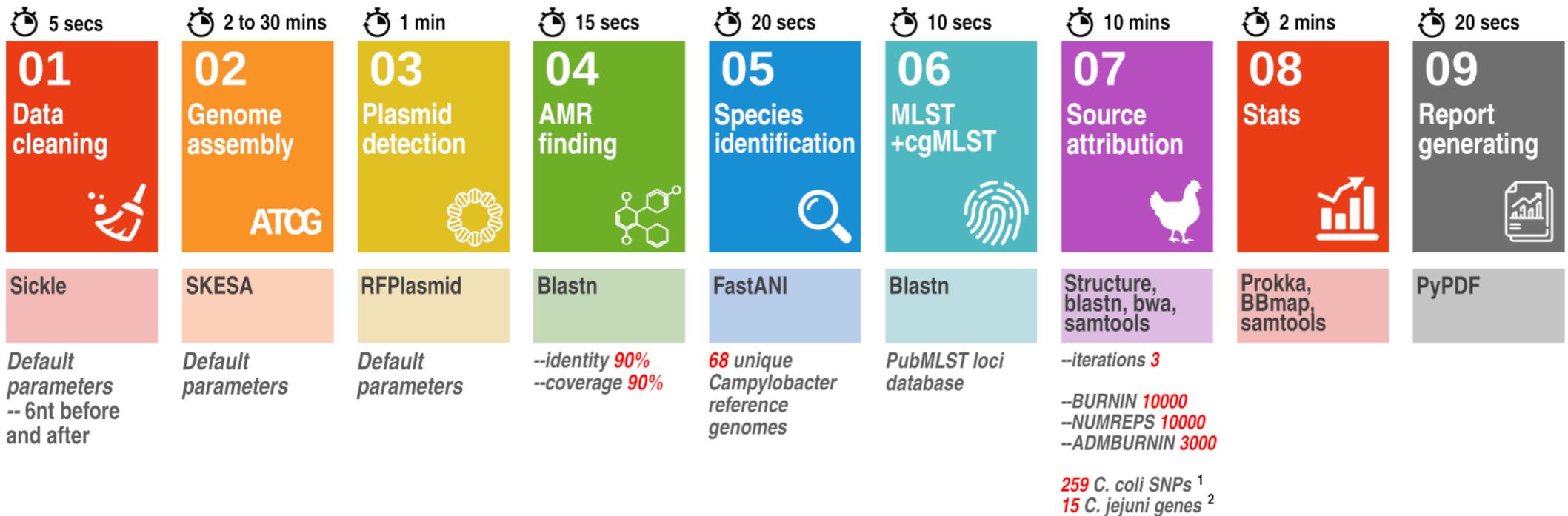
- La source potentielle de contamination est estimée à partir des données publiées par le CNRCH, en collaboration avec l'Université de Bath. <sup>1</sup>Thépault et al. 2017 (*C. jejuni*), <sup>2</sup>Jehanne et al. 2020 (*C. coli*).

# Pipeline d'analyses bio-informatiques



→ Détermination de la taille du génome, du nombre de CDS, du GC% et de la couverture.

# Pipeline d'analyses bio-informatiques



→ Les résultats sont condensés dans un document .pdf.

# Rapport en .pdf contenant les informations essentielles



**CENTRE  
HOSPITALIER  
UNIVERSITAIRE  
BORDEAUX**

Centre National de Référence des Campylobacters & Helicobacters  
Laboratoire de Bactériologie  
CHU Pellegrin  
Place Amélie Raba Léon  
33079 Bordeaux cedex  
Tél: +33 (0)5 58 79 59 77  
Fax: +33 (0)5 57 82 19 77



Informations patient		Informations sur le séquençage	
Nom	██████████	Numéro de l'échantillon	231083491601
Prénom	██████████		
Date de naissance	09/11/1980	Date de prélèvement	18/08/2023
Age	42	Nature	Selles
Sexe	Féminin		
		Nom du run	0025-INTEGRAGEN-2023-P2
		Génome n°	01560
		Projet	CAMPYLOBACTERS ERY-R
		Mois	NOVEMBRE
		Année	2023

Identification de l'espèce, typage moléculaire et attribution de la source de contamination		
Espèce identifiée par ANI (référence)	Complexe clonal (CC)	828
<i>CAMPYLOBACTER COLI LMG6440</i>	Sequence type (ST)	828
	Core genome ST (cgST)	INCONNU (31.348%)
Score ANI (significativité ≥ 95%)	<i>aspA ; glnA ; gItA ; gIyA ; pgm ; tkt ; uncA</i>	Score d'attribution (significativité ≥ 70%)
98.924 %	33 ; 39 ; 30 ; 82 ; 104 ; 43 ; 17	100.0 %

Détermination moléculaire de la résistance aux antibiotiques		
<b>CIP</b> GyrA T86I Contig 1 - 99.9% IDT - 98.1% COV Positions [ 219096 ; 221612 ]	<b>SPC</b> ANT(9)-ic-aad9 Contig 16 - 99.9% IDT - 100.0% COV Positions [ 5880 ; 6656 ]	
<b>ERY</b> erm(B)12 U18931.2 Contig 16 - 100.0% IDT - 100.0% COV Positions [ 3450 ; 4187 ]		
<b>TET</b> tetO (TRONQUE) - Tetracycline Contig 29 - 99.9% IDT - 42.1% COV Positions [ 1 ; 809 ]		
<b>STR</b> ANT(8)-fhaadE Contig 16 - 100.0% IDT - 100.0% COV Positions [ 246 ; 1112 ]		

Statistiques sur les données de séquençage et le génome assemblé			
Reads bruts (R1)	1,256,271	Taille mappant sur la référence	1,593,137 pb
Reads bruts (R2)	1,256,271	Taille de la référence	1,912,816 pb
Taille totale (R1)	188,440,650 pb	Génome couvert à	83.288 %
Taille totale (R2)	188,440,650 pb	Profondeur moyenne	83.917
Reads nettoyés (R1)	1,221,639	Taille du génome assemblé	1,687,533 pb
Reads nettoyés (R2)	1,221,639	Nombre total de contigs	33
Taille nettoyée (R1)	180,342,072 pb	Taille moyenne des contigs	51,137 pb
Taille nettoyée (R2)	176,648,828 pb	GC %	31.42 %
		Nombre de CDS	1,721


# Données patient et run de séquençage

Informations patient				Informations sur le séquençage	
Nom	██████████	Numéro		Nom du run	0025-INTEGRAGEN-2023-P2
Prénom	██████████	de l'échantillon	231083491601	Génome n°	01560
Date de naissance	09/11/1980	Date de prélèvement	18/08/2023	Projet	CAMPYLOBACTERS ERY-R
Age	42	Nature	Selles	Mois	NOVEMBRE
Sexe	Féminin			Année	2023

Identification de l'espèce, typage moléculaire et attribution de la source de contamination		
Espèce identifiée par ANI (référence)		contamination potentielle
CAMPYLOBACTER COLI LMG6440		
Score ANI (significativité ≥ 95%)		tribution (significativité ≥ 70%)
98.924 %	33 ; 39 ; 30 ; 82 ; 104 ; 43 ; 17	100.0 %

**Informations patient**

Cet encart contient à la fois les informations générales concernant le patient, sur la **partie gauche**, mais aussi les informations concernant l'échantillon avec le numéro GLIMS associé (code barre), sur la **partie droite**.

# Identification, MLST et attribution de source

## Identification de l'espèce, typage moléculaire et attribution de la source de contamination

Espèce identifiée par ANI (référence)	Complexe clonal (CC)	828	Source de contamination potentielle
<b>CAMPYLOBACTER COLI CCUG 11283</b>	Sequence type (ST)	899	<b>VOLAILLE</b>
Score ANI (significativité ≥ 95%)	Core genome ST (cgST)	INCONNU (49.367%)	Score d'attribution (significativité ≥ 70%)
98.811 %	<i>aspA ; glnA ; gltA ; glyA ; pgm ; tkt ; unca</i>	33 ; 39 ; 30 ; 82 ; 113 ; 35 ; 17	100.0 %

# Principaux marqueurs de résistance

## Détermination moléculaire de la résistance aux antibiotiques

<b>CIP</b>	GyrA T86I Contig 1 - 99.9% IDT - 98.1% COV Positions [ 219096 : 221612 ]	<b>SPC</b>	ANT(9)-Ic-aad9 Contig 16 - 99.9% IDT - 100.0% COV Positions [ 5880 : 6656 ]		
<b>ERY</b>	erm(B).12_U18931_2 Contig 16 - 100.0% IDT - 100.0% COV Positions [ 3450 : 4187 ]				
<b>TET</b>	tetO (TRONQUÉ) - Tetracycline <b>P</b> Contig 20 - 99.9% IDT - 42.1% COV Positions [ 1 : 809 ]				
<b>STR</b>	ANT(6)-If-aadE Contig 16 - 100.0% IDT - 100.0% COV Positions [ 246 : 1112 ]				

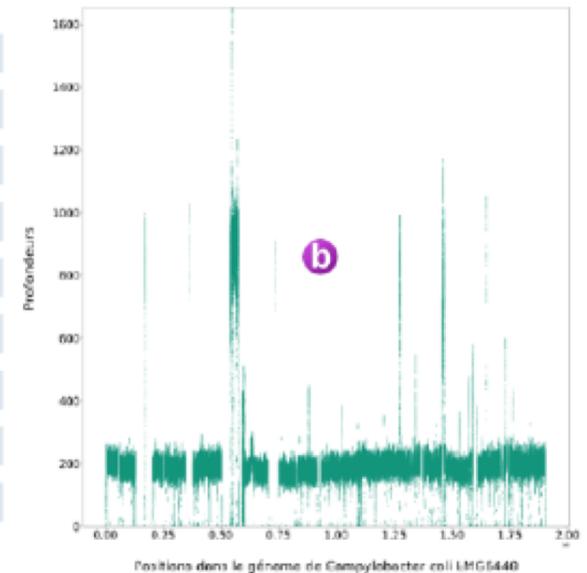
**AMP**=Ampicilline; **CIP**=Ciprofloxacine; **LEV**=Lévofloxacine; **ERY**=Erythromycine;  
**TET**=Tétracycline; **GEN**=Gentamicine; **AMC**=Amoxicilline-Acide Clavulanique;  
**CLA**=Clarithromycine; **AMX**=Amoxicilline; **STR**=Streptomycine; **KAN**=Kanamycine;  
**SPC**=Spectinomycine; **CAP**=Chloramphenicol; **MET**=Métronidazole; **RIF**=Rifampicine;  
**LIN**=Lincosamide.

**P** Identification d'un mécanisme de résistance porté par un plasmide (outil **RFPlasmid**).

# Informations sur la qualité de séquençage

## Statistiques sur les données de séquençage et le génome assemblé

<b>Reads bruts (R1)</b>	1,256,271	<b>Taille mappant sur la référence</b>	1,593,137 pb
<b>Reads bruts (R2)</b>	1,256,271	<b>Taille de la référence</b>	1,912,816 pb
<b>Taille totale (R1)</b>	188,440,650 pb	<b>Génome couvert à</b>	83.288 %
<b>Taille totale (R2)</b>	188,440,650 pb	<b>Profondeur moyenne</b>	83.917
<b>Reads nettoyés (R1)</b>	1,221,639	<b>Taille du génome assemblé</b>	1,687,533 pb
<b>Reads nettoyés (R2)</b>	1,221,639	<b>Nombre total de contigs</b>	33
<b>Taille nettoyée (R1)</b>	180,342,072 pb	<b>Taille moyenne des contigs</b>	51,137 pb
<b>Taille nettoyée (R2)</b>	176,648,828 pb	<b>GC %</b>	31.42 %
		<b>Nombre de CDS</b>	1,721



- a** La taille des données brutes (fastq) après séquençage, avant nettoyage par Sickle (partie haute) et après nettoyage (partie basse).
- b** La couverture sur l'espèce identifiée (ou le pourcentage des séquences obtenues qui correspond à cette espèce) et la profondeur moyenne de séquençage c-a-d le nombre de fois qu'une séquence est obtenue (qualité minimale acceptable autour de 30X).
  - c** Le nombre de contigs (=séquences) obtenu après assemblage du génome (nombre maximum acceptable = 100) et la taille moyenne de ces séquences (plus la taille est importante, plus la qualité du génome est meilleure).
  - d** GC% = Proportion de nucléotides étant un G ou un C (taux spécifique à l'espèce); CDS=Coding Sequence, soit le nombre de gènes au sein du génome.

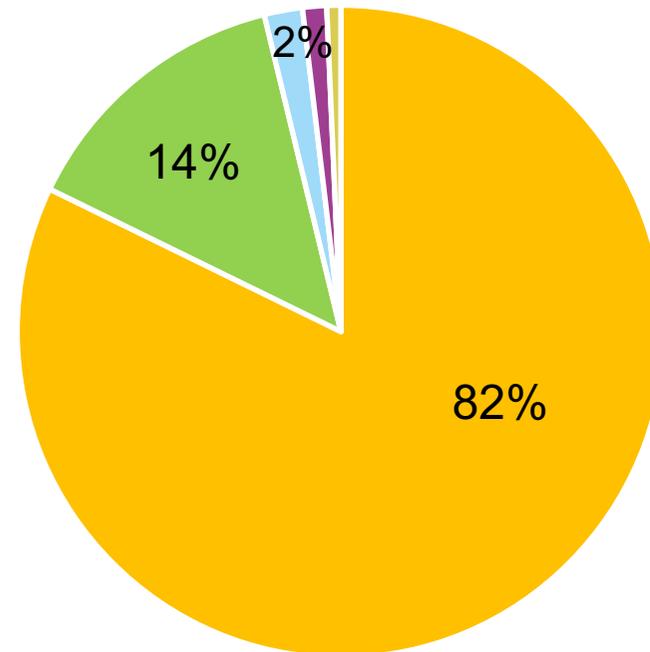


## Bilan de février à juillet 2024



## Bilan février - fin juillet 2024

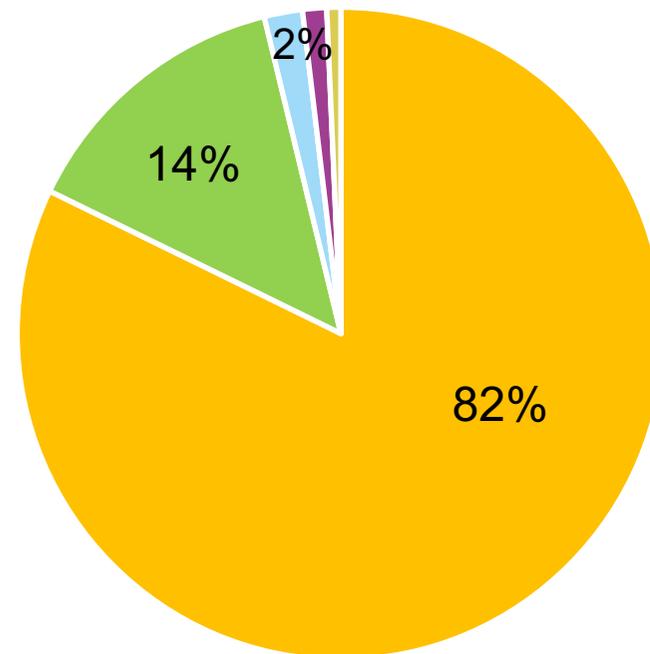
- Nombre de génomes séquencés
  - › 1310 génomes soit 218/mois
- Données épidémiologiques
  - › 758 hommes, 553 femmes : sex ratio = 1,4
  - › 1248 selles (95,3%), 54 hémocultures (4,1%)



■ *C. jejuni*   ■ *C. coli*   ■ *C. fetus*  
■ *A. butzleri*   ■ Autres

## Bilan février - fin juillet 2024

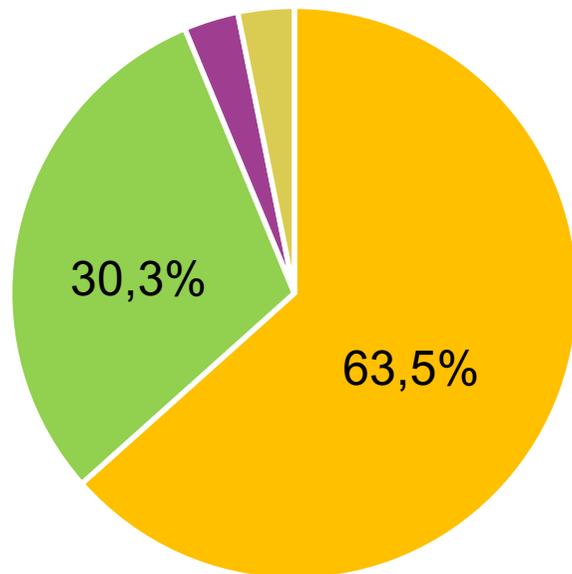
- Nombre de génomes séquencés
  - › 1310 génomes soit 218/mois
- Données épidémiologiques
  - › 758 hommes, 553 femmes : sex ratio = 1,4
  - › 1248 selles (95,3%), 54 hémocultures (4,1%)
- Concordance identification
  - › rares discordances : 23/1310 (1,8%)
  - › au genre (n=3, *A. butzleri*)
  - › à l'espèce (n=20)
  - › aucune discordance MALDI\_CNRCH versus NGS



■ *C. jejuni* ■ *C. coli* ■ *C. fetus*  
■ *A. butzleri* ■ Autres

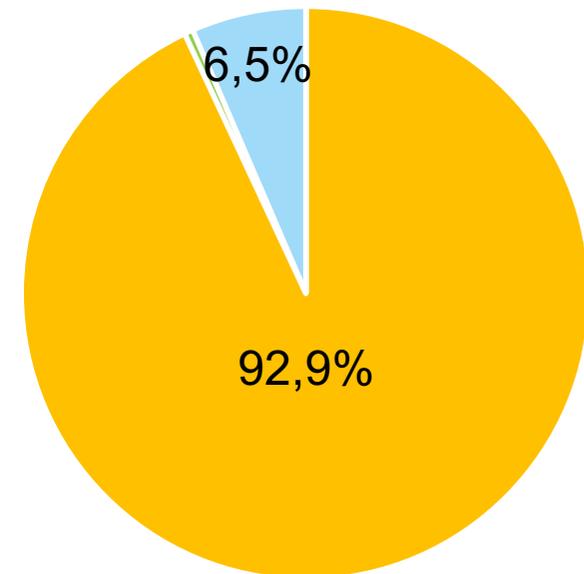
## Attribution de sources par NGS

*C. jejuni*



- Poulet
- Ruminant
- Porc
- Environnement
- Non identifié

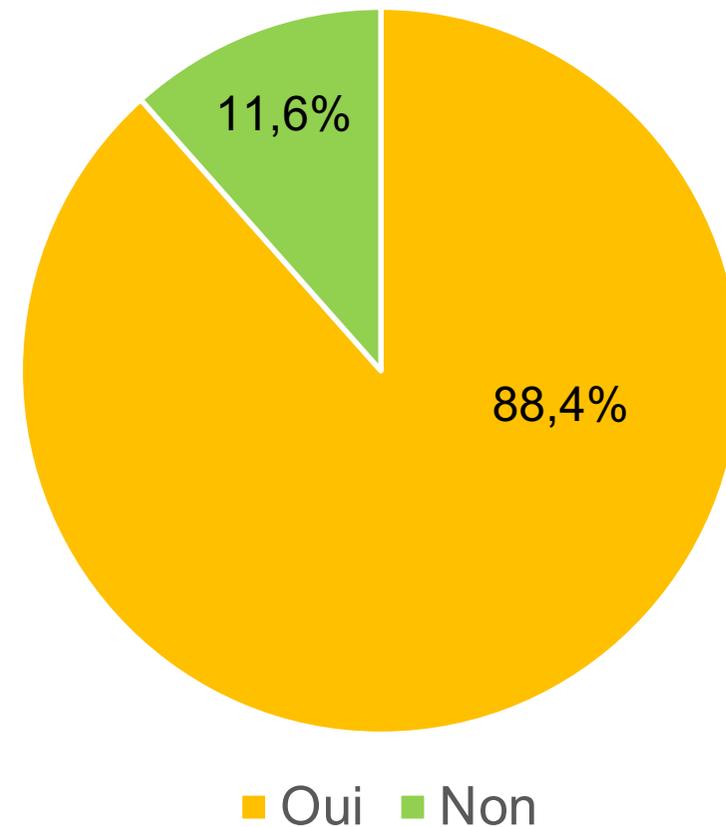
*C. coli*



- Poulet
- Ruminant
- Porc
- Environnement
- Non identifié

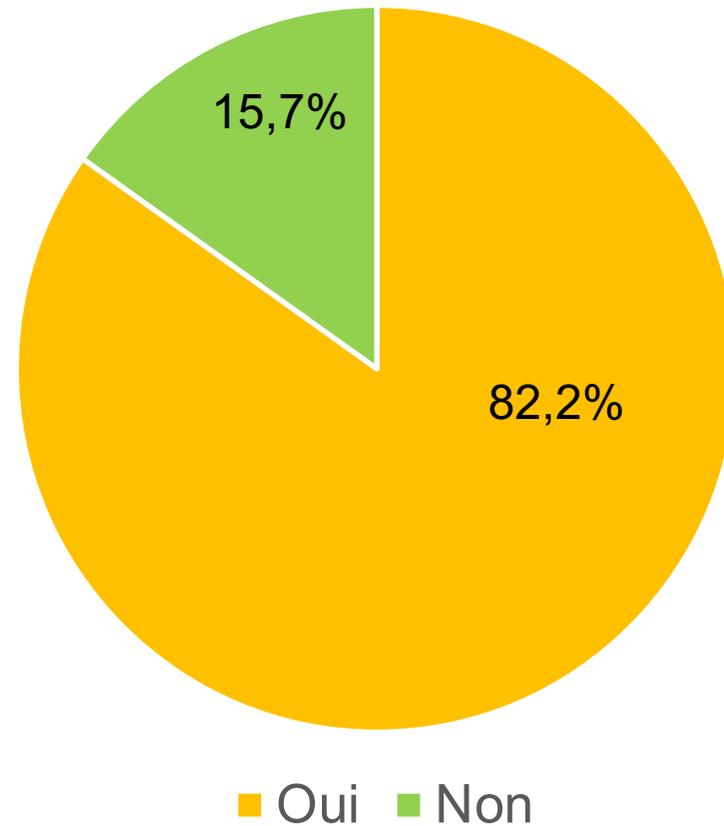
## Concordance phénotype/résistome pour *C. jejuni*

- Excellents résultats
- Principales discordances
  - › ampicilline (42,4%)
  - › ciprofloxacine (27,2%)
  - › tétracycline (27,2%)



## Concordance phénotype/résistome pour *C. coli*

- Excellents résultats
- Principales discordances
  - › ampicilline (34,5%)
  - › érythromycine (17,2%)
  - › tétracycline (13,8%)



## Suivi des résistances par l'étude du résistome

	<i>C. jejuni</i> (n=1076)	<i>C. coli</i> (n= 184)	Mécanismes principaux
Ampicilline	29,2%	25,9%	prom. <i>blaOXA61</i>
Ciprofloxacine	57,1%	63,6%	<i>gyrA</i> (T86I)
Erythromycine	0,74%	12,5%	ADNr 23S + <i>ermN</i>
Tétracycline	35,4%	81,5%	<i>tet0</i>
Gentamicine	0,28%	2,2%	<i>aph2''</i>
Kanamycine	1,3%	5,4%	<i>aph3''-IIIa</i>
Streptomycine	3,4%	31,5%	<i>ant6, spw</i>
Spectinomycine	0,28%	3,3%	<i>ant9</i>

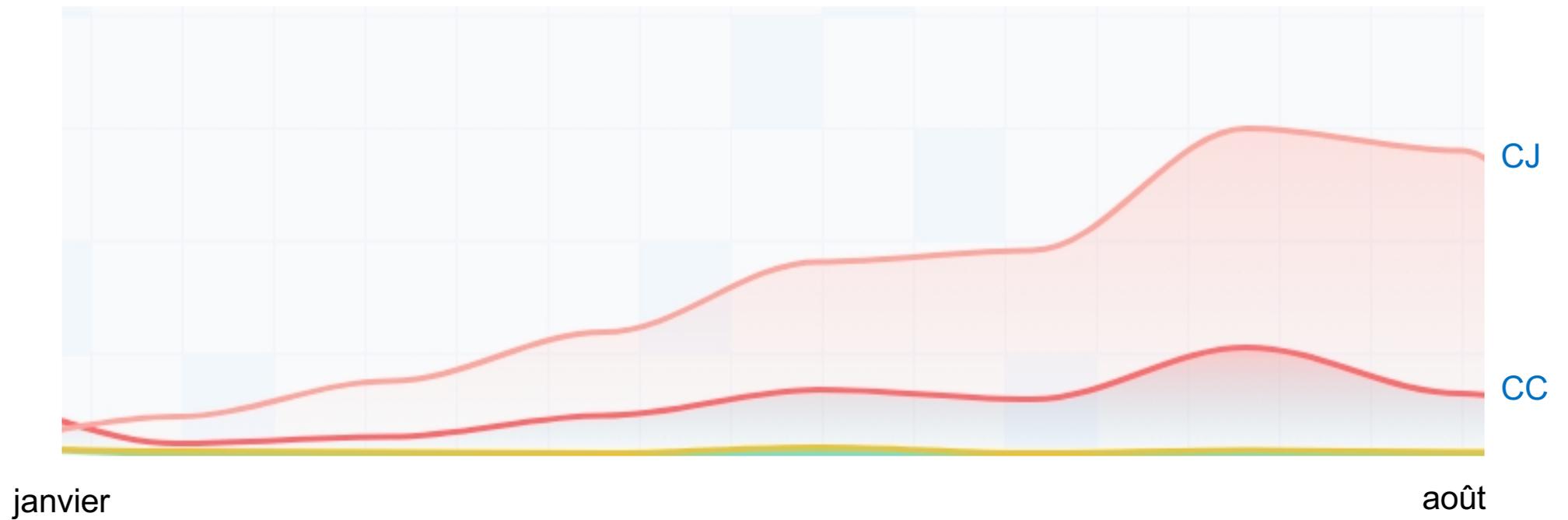
## Suivi des résistances par l'étude du résistome-comparaison avec 2023

	NGS-2024	2023	NGS-2024	2023
	<i>C. jejuni</i> (n=1076)	<i>C. jejuni</i> (n=6484)	<i>C. coli</i> (n= 184)	<i>C. coli</i> (n= 995)
Ampicilline	29,2%	34%	25,9%	28,4%
Ciprofloxacine	57,1%	64,8%	63,6%	66,2%
Erythromycine	0,74%	0,5%	12,5%	7,2%
Tétracycline	35,4%	44,1%	81,5%	78,9%
Gentamicine	0,28%	0,4%	2,2%	3%



# Bilan général au 18 septembre 2024

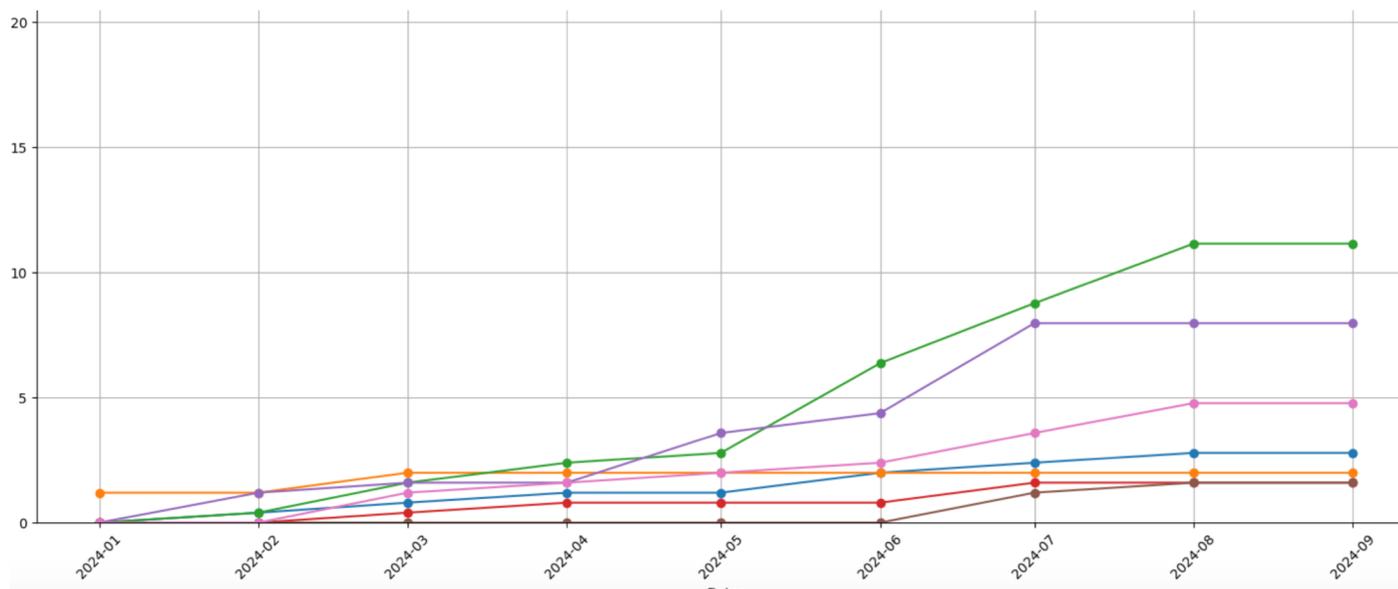
→ NGS sur 41% des souches reçues au CNR



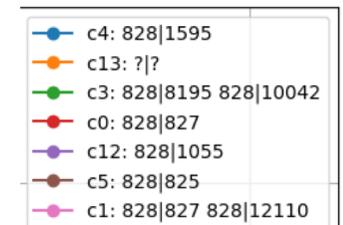
# Suivi en France des clones circulants par core-genome MLST

## → Analyse du cgMLST

- › augmentation depuis juin d'un clone en France de *C. coli* : CC 828, ST 8195 (30 souches)



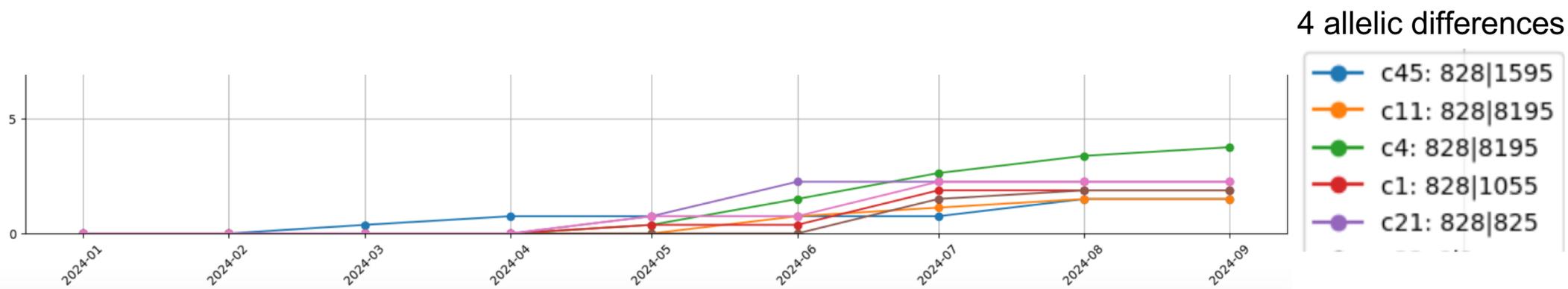
50 allelic differences



# Suivi en France des clones circulants par core-genome MLST

## → Analyse du cgMLST

- › augmentation depuis juin d'un clone en France de *C. coli* : CC 828, ST 8195 (10 souches)



## → Possibilité de comparaison à des données européennes

- › alerte en août 2024 en Irlande non confirmée en France : *C. jejuni* CC 353-ST 8334

# Conclusions

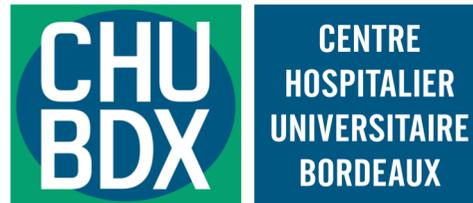
- Outils performants pour le suivi épidémiologique des souches de *Campylobacter*
  - identification de genre et espèce
  - attribution de source
  - classification MLST (ST, CC)
  - marqueurs génotypiques de résistance
  - typage de cas groupés d'infections
  
- Excellente concordance résistome/phénotype
  - suivi des nouveaux mécanismes de résistance (nouveau promoteur de la *blaOXA*)
  
- Comparatifs nationaux, internationaux et One Health

Merci pour vos envois de souches au CNRCH

Vous êtes formidables!



Merci pour votre attention !



**Directeur** : Philippe Lehours

**Biologiste** : Marine Jauvain

**Ingénieurs** : Lucie Bruhl-Bénéjat, Quentin Jehanne, Léo Gillet

**Techniciennes** : Astrid Ducournau, Johanna Aptel, Claudie Perreau, Marie Taymont

**Secrétaire** : Erick Keisler

