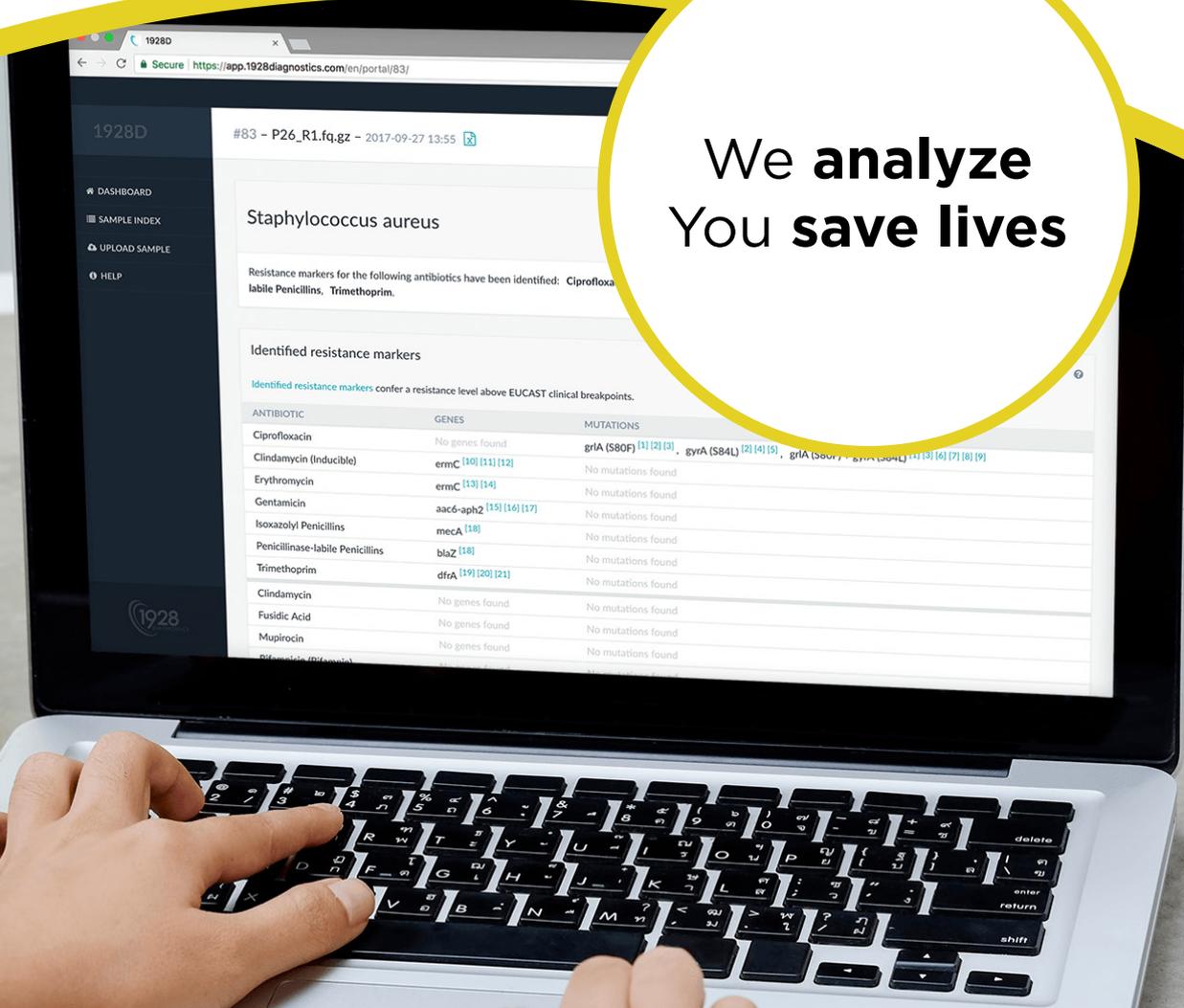




Outbreak analysis of *Shigella sonnei* using 1928 SNP analysis

A white paper describing how to use the 1928 platform for outbreak analysis

We analyze
You save lives



Introduction

1928 Diagnostics offers a platform for analysis of NGS sequenced bacterial genome data for resistance predictions and outbreak surveillance in a clinical routine setting.

For outbreak analysis, two different methods can be used: cgMLST and SNP analysis.

For cgMLST, 1928 Diagnostics support a wide range of bacteria through predefined and manually curated schemas. In this whitepaper, we are looking at data from a previously published study and re-create the phylogeny presented in the article, by performing SNP analysis.

Article and raw data

The *Shigella sonnei* outbreak in focus in this whitepaper is:

First report of sexually transmitted multi-drug resistant *Shigella sonnei* infections in Switzerland, investigated by whole genome sequencing
Hinic Vladimira, Seth-Smith Helena, Stöckle Marcel, Goldenberger Daniel, Egli Adrian

doi:10.4414/smw.2018.14645
Swiss Med Wkly. 2018;148:w14645

The article describes the first report of 3 cases of sexually acquired *Shigella sonnei* infections in Switzerland. The data from the article can be found in the Bioprojects PRJEB23646 and PRJNA315192.

Data analysis

The raw FASTQ files for each sample from the Bioprojects were uploaded to the 1928 platform using the *Escherichia coli* pipeline. The analysis starts automatically once the upload is complete and will perform de-novo assembly as well as identifying resistance markers.

Upload new samples

Upload samples to have them automatically analyzed. We currently support paired-end Illumina and single-end Ion Torrent reads. Read more in our [help page](#).

Select format

Illumina (paired end) ▾

Select pipeline

Escherichia coli ▾

Add files

Drop files here or click to browse.

Use gzipped fastq format (.fq.gz or .fastq.gz files).

⬆️ UPLOAD SAMPLES

Once the analysis is done, the sample index shows an overview:

ID	NAME	TOXINS	MLST	PHYLOGROUP	DATE	STATUS
71	SRR3530636	■■■	152		2019-09-11 13:15	Done
70	SRR3530555	■■■	152		2019-09-11 13:15	Done
69	SRR3530549	■■■	152		2019-09-11 13:15	Done
68	SRR3530542	■■■	152		2019-09-11 13:15	Done
67	SRR3530516	■■■	152		2019-09-11 13:15	Done
66	SRR3530452	■■■	152		2019-09-11 13:15	Done
65	SRR3530435	■■■	152		2019-09-11 13:15	Done
64	SRR3530422	■■■	152		2019-09-11 13:15	Error
63	SRR3530393	■■■	152		2019-09-11 13:15	Done

One of the samples failed to pass the 1928 Quality Control with an estimated coverage depth of 24x and was subsequently

excluded from further analysis. This sample is shown in red in the screenshot above, sample ID SRR3530422.

To perform the SNP analysis - the sample set is selected, and the SNP button is pressed.

Overview - Escherichia coli

cgMLST | SNP | Export | SEARCH | RELOAD

ID	NAME	TOXINS	MLST	PHYLOGROUP	DATE	STATUS
<input checked="" type="checkbox"/> 92	SRR5029673	■■■	152		2019-09-11 14:01	Done
<input checked="" type="checkbox"/> 91	SRR5006072	■■■	152		2019-09-11 14:01	Done
<input checked="" type="checkbox"/> 90	SRR5005299	■■■	152		2019-09-11 14:01	Done
<input checked="" type="checkbox"/> 89	SRR4788222	■■■	152		2019-09-11 14:01	Done
<input checked="" type="checkbox"/> 88	SRR4787904	■■■	152		2019-09-11 13:15	Done
<input checked="" type="checkbox"/> 87	SRR4787219	■■■	152		2019-09-11 13:15	Done
<input checked="" type="checkbox"/> 86	SRR4786326	■■■	152		2019-09-11 13:15	Done
<input checked="" type="checkbox"/> 85	SRR4195740	■■■	152		2019-09-11 13:15	Done

Pick reference

ACCESSION	TITLE	NCBI
<input checked="" type="radio"/> CP000038.1	Shigella sonnei Ss046, complete genome	Link
<input type="radio"/> Add new reference		

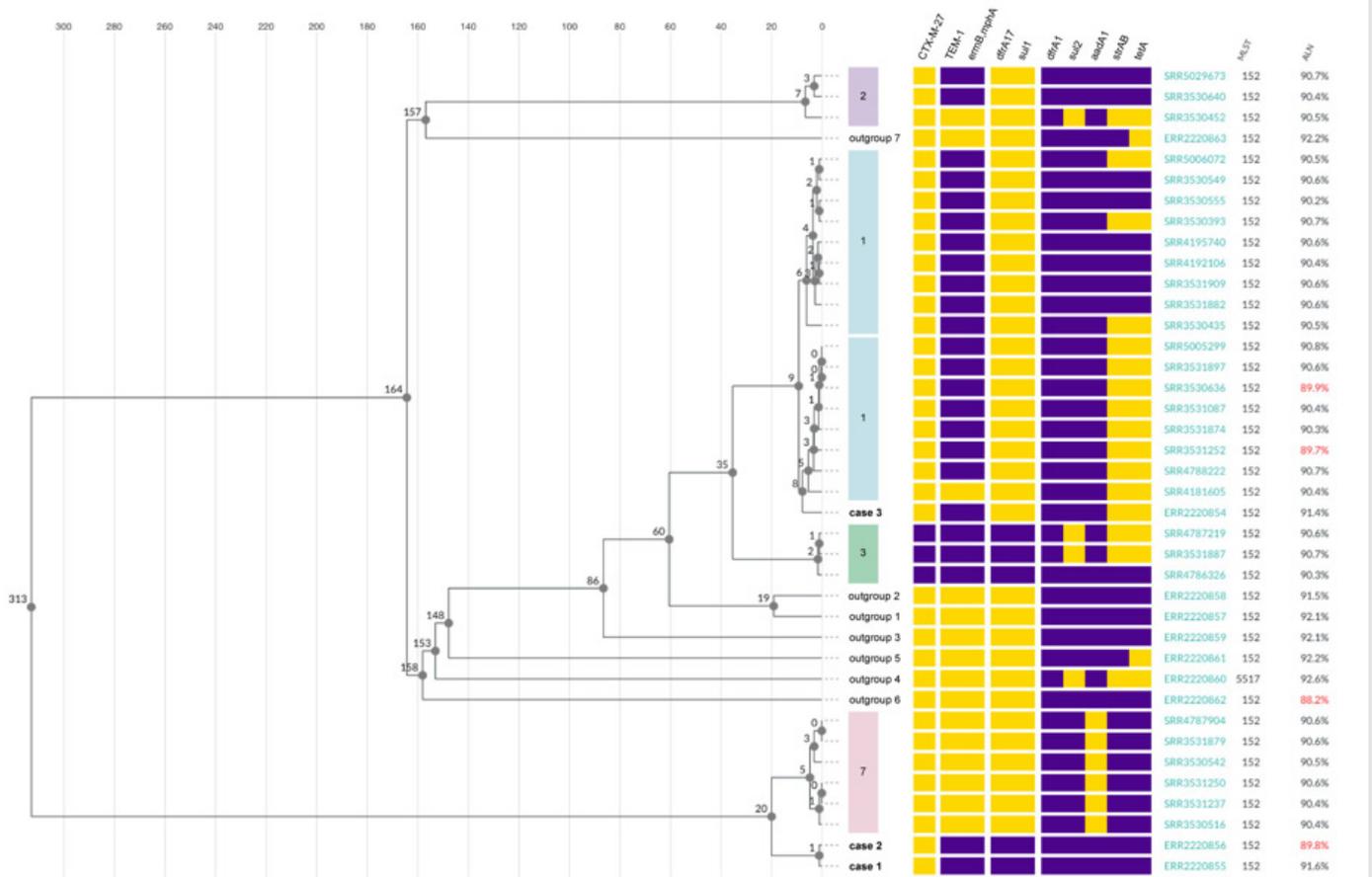
CLUSTER

The SNP analysis pipeline is based on read alignment against a reference, and in the analysis, the same reference genome was selected as in the original article, namely the Shigella sonnei Ss064, NCBI accession number CP000038. The dendrogram tree generated from the SNP output is created from the distance matrix using the UPGMA algorithm.

In addition, the resistance genes found in the samples were exported to an Excel document, where the presence of genes were combined into a presence/absence table and overlaid on the dendrogram, together with the UK MSM cluster numbering.

Results

The 1928 SNP analysis is able to re-create the phylogeny from the original article with great detail. The topology of the original tree is preserved, including the two cluster groups within the UK MSM cluster 1.



Conclusions

SNP analysis is a powerful tool for outbreak surveillance to use as a complement to cgMLST, for pathogens where no cgMLST schemas exist, or where additional resolution might be needed.

To learn more about the 1928 platform visit

www.1928diagnostics.com

For resources about using the platform and other articles:
https://1928diagnostics.com/product_resources/index.html

Scientific support

Need to complete your research with a nice-looking typing schema or an AMR profiling? 1928 offers to open the platform for research samples. Just connect with us at:

support@1928diagnostics.com

Schedule a meeting or book a demo?
Email us at: **sales@1928diagnostics.com**

1928 Diagnostics
Stena Center 1B, 412 92
Gothenburg
Sweden



www.1928diagnostics.com