IR Biotyper®

Ready-to-use *E. coli* STEC Classifier

Identification of Shiga toxin-producing *Escherichia coli* (STEC) O-serogroups based on machine learning

Shiga toxin-producing *Escherichia coli* (STEC) are important foodborne pathogens capable of causing diseases ranging from mild diarrhea to severe complications such as hemorrhagic colitis and hemolytic uremic syndrome (HUS). Rapid identification of STEC is critical for outbreak investigations, food safety monitoring, and public health surveillance. Conventional STEC detection typically relies on molecular assays often requiring additional laboratory resources. Therefore, complementary tools that support efficient screening and classification workflows will benefit those laboratories.

Fourier-transform infrared (FTIR) spectroscopy, combined with the new ML-based *E. coli* STEC Classifier for the IR Biotyper, provides a rapid and standardized approach to identify the O-serogroups from STEC isolates. The classifier leverages unique infrared fingerprint profiles and advanced machine learning algorithms to support routine laboratory workflows, helping reduce hands-on time while improving operational efficiency.

Benefits of using Classifiers:

- Ready-to-use
- Immediate and clear result display by color coding and score values
- Easy and fast workflow improving lab efficiency
- Transparent, objective and standardized serogrouping results
- Reduces complexity of routine characterization workflows

What's a Classifier?

An FTIR classifier is a software tool that learns from infrared spectra of known training classes and uses AI-based algorithms to confidently classify new, unknown spectra (Fig.1).

How was the *E. coli* STEC Classifier developed?

The *E. coli* STEC classifier was trained to identify isolates of the species *E. coli* belonging to O-serogroups commonly associated with Shiga-toxin producers (STEC). The covered O-serogroups are O26, O91, O103, O104, O111, O145, and O157.

To develop the Classifier, 294 isolates were included in the training set, 61 of which belonged to STEC-associated O-serogroups. In total, 3,015 spectra were used to train a support vector machine (SVM) algorithm within the IR Biotyper software. After training, the classifier was evaluated using a set of 10,250 spectra from 1,970 isolates that were not included in the previous training dataset. The accuracy of the unknown classification is shown below (Table 1).

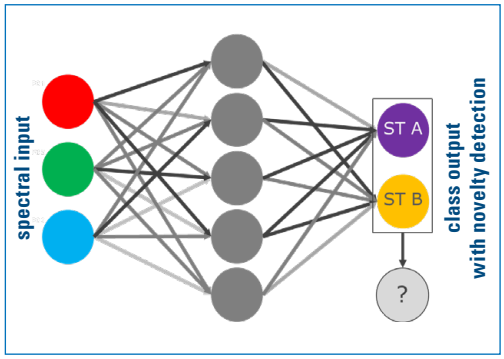

Fig. 1 Schematic depiction of a Classifier that can be applied to unknown samples for their classification (right side). Classifier creation involves a training set of isolates (left side) to establish a recognition network (middle).

		Prediction									
		O104	O111	O145	O157	O26	O91	negative	uncertain	∑ Actual	Class recall
Actual	O104	3	0	0	0	0	0	0	0	3	100%
	O111	0	1	0	0	0	0	0	0	1	100%
	O145	0	0	7	0	0	0	0	0	7	100%
	O157	0	0	0	18	0	0	0	0	18	100%
	O26	0	0	0	0	18	0	0	0	18	100%
	O91	0	0	0	0	0	1	0	0	1	100%
	negative	1	0	0	1	1	0	1915	4	1922	99.6%
	∑ Prediction	4	1	7	19	19	1	1915	4	1970	
	Class precision	75%	100%	100%	95%	95%	100%	100%			

Table 1. Confusion matrix of the test results.

No. 1931332 © 08-2026 Bruker Daltonics

Not for use in clinical diagnostic procedures.
Please contact your local representative for availability in your country.

IR Biotyper® is a registered trademark of the Bruker group of companies.

Online information
bruker.com/microbiology

Bruker Daltonics GmbH and Co. KG
Bremen · Germany
Phone +49 (0) 421-2205-0

info.md@bruker.com

Bruker Scientific LLC
Billerica, MA · USA
Phone +1 (978) 663-3660