ThermoFisher SCIENTIFIC

The world leader in serving science

May 26, 2020

TECHNICAL NOTE

Artifacts Identified Post-Developmental Validation: Yfiler Plus™ PCR Amplification Kit

Revision History

Revision	Date	Description
A	May 2020	Initial publication

The purpose of this document is to assist with data interpretation by providing a repository of artifacts, identified and characterized as a result of investigating customer reports following developmental validation of the Yfiler Plus kit (PN 4482730 and 4484678). Many of the artifacts included in this document fall below the peak amplitude threshold (PAT) used during developmental validation, appear outside of the read region, or are attributable to specific sample types that were not encountered during developmental validation.

Background

During developmental validation of the Yfiler Plus PCR Amplification Kit, a 175 RFU PAT was applied to data generated on the 3500/xL Genetic Analyzer. Many laboratories, following internal Yfiler Plus validations, implemented PATs below 175 RFU. As a result of lower analysis thresholds, low level artifacts not identified in the developmental validation have been reported.

Due to the wide variety of sample and substrate types processed in forensic laboratories, it is not feasible to process all sample and substrate type combinations during STR kit development and validation testing. Instead, representative samples and substrates are selected and analyzed during this process. Therefore, it is possible for sample-specific artifacts to be identified through customer reports and subsequently tested and characterized in the Thermo Fisher Scientific HID Laboratory. Such artifacts are included in this document to assist with interpretation of data generated from similar sample or substrate types.

A variety of mechanisms, such as dye byproducts, the formation of secondary structures, non-traditional stutter, non-specific binding or non-human interaction can introduce artifacts into PCR STR data. The detectable presence and intensity (peak height) of these types of artifacts can be

dependent upon amplification and/or electrophoresis conditions, such as elevated PCR cycle numbers or CE system sensitivity differences.

Method

Artifacts detailed in this technical note were detected and characterized post-product development. Artifacts that are the result of standard PCR phenomena, such as traditional stutter or minus A peaks, or that were observed during the species specificity studies of the developmental validation are not included. See the Yfiler Plus kit user guide (publication number 4485610) for information on these types of artifacts.

The Summary Table lists reproducible DNA-dependent artifacts that are included and defined in the Yfiler Plus User Guide. For this reason, these artifacts are not further described in the Detailed Artifact Descriptions section of this document. Instead, see the Yfiler Plus kit user guide for more information on these specific artifacts.

The data herein is specific to Yfiler Plus kit data run on a 3500 series instrument, and unless otherwise noted, was generated following all manufacturer recommended PCR and electrophoresis conditions, as documented in the Yfiler Plus kit user guide. The information provided also applies to Yfiler Plus data run on other electrophoresis platforms, although peak heights and base pair sizes may differ.

Newly identified and characterized artifacts will be periodically added to this document as applicable.

Summary Table: Click on the hyperlinks for more information for artifacts that are not already detailed in the Yfiler Plus User Guide.

Dye Channel	Locus	Artifact ID	Approximate Base Pair (bp) Size
6-FAM™	**DYS389II**	YFP_DYS389II-1	270-271 bp
		YFP_DYS389II-2	280-281 bp
	DYS627	YFP_DYS627-1	348-349 bp
	DYS576	YFP_DYS576-1	116 bp
VIC™	**DYS391**	YFP_DYS391-1	N-10 bp
		YFP_DYS391-2	N-5 bp
	N/A	YFP_VIC-1	70 bp**
NED™	**N/A**	YFP_NED-1	441 bp**
TAZ™	**DYS570**	YFP_DYS570-1	139-140 bp
		YFP_DYS570-2	144-145 bp
	DYS437	YFP_DYS437-1*	N-5 bp
		YFP_DYS437-2*	N-12 bp
		YFP_DYS437-3*	N-16 bp
	DYS385	YFP_DYS385-1	225-260 bp
	N/A	YFP_TAZ-1	412-413 bp**

*These artifacts were detected and characterized post-product development but are included in the most recent User Guide revision.

**These artifacts appear outside the read region.

Detailed Artifact Descriptions

Dye Channel: 6-FAM

Locus: DYS576

Artifact ID: YFP_DYS576-1

- Location: ~116 bp
- Probable cause: sample dependent artifact likely caused by the presence and amplification of non-human DNA in the sample
- Typical signal intensity: relative to the amount of non-human DNA present
- Additional Info:
    - Has been observed in a variety of crime scene related sample types (for example, swabs of a door handle and T-shirt, fingernail scrapings, and a post-coital sample).
    - Sequencing studies have shown that the artifact sequence has partial homology with sequences from the *Comamonas* spp. and *Acidovorax* spp. genomes, both members of the *Comamondacea* family. At the time of the search, the NCBI BLAST statistics for these partial matches ranged from a low E value of 5.0e-08 with 51% query coverage and 91.49% identity for *Comamonas thiooxydans* to a high E value of 3.0e-04 with 38% query coverage and 94.29% identity for *Acidovorax* sp. T1. Based on the results of the BLAST search, the artifact sequence has no homology to the human genome.[1]

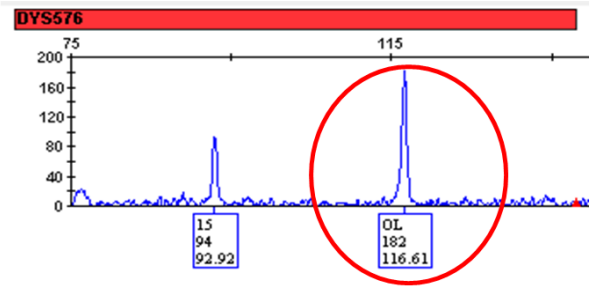

Figure 1: Sample dependent artifact (YFP_DYS576-1), obtained from a crime scene sample, at 116.61 bp in DYS576. The y-axis is scaled to 200 RFU, and the artifact is 182 RFU.

Dye Channel: VIC

Artifact ID: YFP_DYS391-2

- Location: N-5 bp (typically observed ~5 nucleotides before the parent DYS391 peak)
- Probable cause: DNA template dependent artifact likely caused by polymerase slippage (atypical stutter) due to a repeating nucleotide region within the target sequence range
- Typical signal intensity: relative to DNA input amount (<2% of the parent peak height)[2]
- Additional Info:
    - The same repeating region and related artifact is also present at the DYS391 marker in the GlobalFiler™ PCR Amplification Kit.

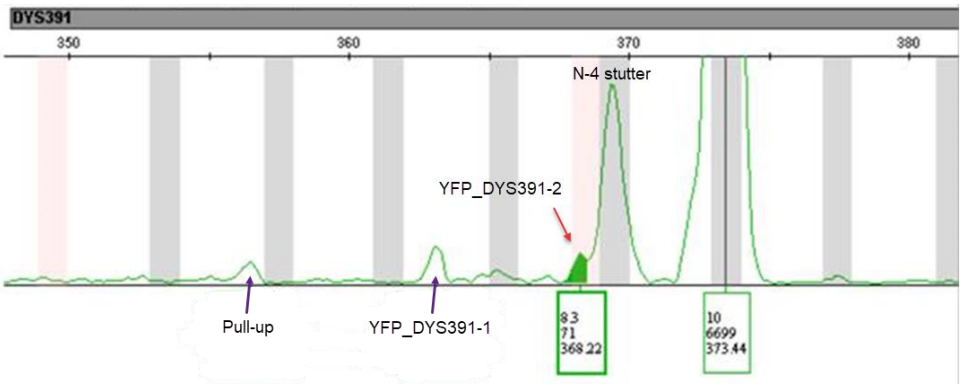

Figure 2: PCR artifact (YFP_DYS391-2) located ~5 bp before the 10 allele as indicated by the red arrow. The y-axis is scaled to 100 RFU, and the artifact is 71 RFU, which is ~1.2% of the parent 10 allele. This example also displays the N-10 bp (YFP_DYS391-1) artifact that is described in the Yfiler Plus User Guide.

Dye Channel: NED

Locus: N/A

Artifact ID: YFP_NED-1

- Location: ~441 bp
- Probable cause: DNA template dependent artifact likely caused by non-specific STR amplification in male samples
- Typical signal intensity: relative to DNA input amount
- Additional info:
    - Because this artifact is far outside the read region, it should not interfere with interpretation - further characterization was not completed

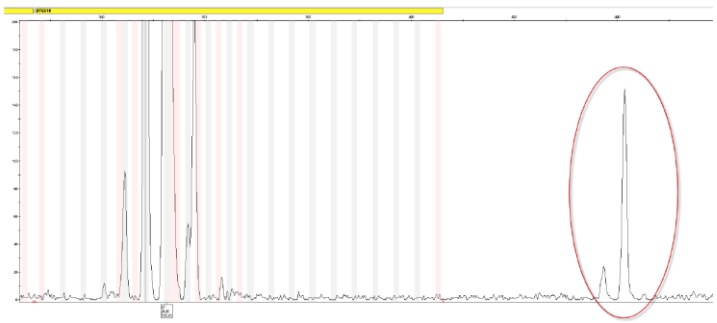

Figure 3: PCR artifact (YFP_NED-1), obtained from 1 ng 007 control DNA, located outside of the DYS518 marker range. The y-axis is scaled to 200 RFU.

Conclusions

Thermo Fisher Scientific's Human Identity STR kits are subjected to rigorous development specifications and criteria to prevent the presence of artifacts. However, the diversity of sample types, variation in storage conditions, and the use of analysis approaches that promote higher sensitivity for the forensic application can lead to the observation of artifacts not typically observed during development. This document will be periodically updated with additional reported and characterized artifacts identified post-product development to assist in the analysis and interpretation of Yfiler Plus STR results.

Comments

[1]Refer to https://blast.ncbi.nlm.nih.gov/Blast.cgi for information on BLAST statistics. BLAST statistics and matches observed can change over time as new sequences are added to the database. Match information and the associated statistics are provided for informational purposes only and do not suggest that a particular artifact was conclusively caused by the named organism.

[2]The observed percentages are provided for guidance only and are based on available data. These percentages have not been confirmed to the same extent as, for example, average marker stutter percentages.