BAX® System Real-Time PCR Algorithm Enhancements

In recent years Hygiena® has enhanced and optimized the determination algorithms for the BAX® System assays. Ascertaining the presence of microorganisms in a sample by analyzing the data output of a PCR instrument is a highly complex biochemical and computational task. Current PCR technologies, both endpoint (melt) and real-time (RT), rely on the reactions of fluorescent dyes to detect the amplification of targeted DNA. The BAX Q7 System instruments use an array of filters, lenses, prisms and sensors to measure the luminescence generated by these reactions. By analyzing the luminescence signals captured by the sensors, we can determine if a pathogen is present in a sample.

Not every food matrix or environmental condition can be included in initial algorithm development. Therefore, Hygiena has endeavored to develop more robust and flexible algorithms. Each BAX System assay may use different algorithm types. This document summarizes the differences across the portfolio.

Model description and nomenclature

The current RT algorithms for the BAX System Q7 have been improved and modernized in the past several years. BAX System Q7 software version 4.0, our first Windows 10-only software version, introduced a modern MATLAB engine for data analysis and a new SDS equipment driver, which enabled Hygiena to reduce errors and improve the operation of the Q7 instrument and software. These changes also allowed Hygiena to utilize modern statistical algorithm development tools better. Extensive validations have been performed across all assays to prove that there were no changes in critical results or analysis.

Each type of algorithm model may be structured and trained according to the quality and availability of the data for each assay, as well as the type of assay (i.e., melt vs. real-time). Below is a classification of the current BAX System Q7 algorithms:

1. **Assay-Specific Models:** In this category, each assay has its own set of analysis functions which were developed using only data generated from each individual assay. Also, the models are trained for each component of the assay, e.g., the model for the target of an assay is different from the model for the internal PCR control (IPC) of the same assay. Currently, there are two types of assay-specific algorithms:

    a. **Frequentist Statistical Models**: These historical models were developed from extensive statistical analysis of assay-specific PCR data to create thresholds to make positive and negative result interpretations. They usually consist of an accurate, although slow curve-fitting routine, followed by a classification routine. We've found that the optimization routines in these legacy curve fitting functions might be susceptible to differences between platforms caused by memory allocation. The classification functions of these algorithms consist of estimated rules and thresholds for the curve-fitting parameters to classify individual PCR signals as positive or negative. In general, these models are inflexible and hard to fine-tune when customer support requests are placed for improvement or corrections.

    All BAX System assays currently use this type of model, with the exception of BAX System Real-Time E.coli EXACT and Real-Time Salmonella, which use supervised machine-learning models (since 2020). When and if algorithm refinement is necessary, Hygiena intends to begin our more dynamic, Bayesian statistical-based method, which can quickly learn new situations as new data is presented.

b. **Supervised Machine-Learning Models**
Instead of performing a statistical analysis and setting empiric thresholds for the classification of PCR curves, as done in the statistical models described above, supervised machine-learning models estimate the characteristics of the data and set up "thresholds" on their own. These thresholds are dynamic and consider the interactions between all the features used for classification. Like the old statistical models, these models also rely on the output of curve fitting routines, along with additional features such as signal-to-noise ratio and parameters describing phenomena like exponential decay in signals. The new curve-fitting routines are faster than the previous generation. Because these are completely new curve-fitting functions, there might be differences in the results in Ct and results interpretation when compared to the previous generations of the models.

2. **Unsupervised Machine-Learning Models**
These algorithms are the most recent types of methods we have used for the classification of PCR signals. Unlike the assay-specific models described above, these models use mixed data from different assays and assay components in their training. In this method, there is minimal signal pre-processing, and unlike the assay-specific methods we described above, we do not use curve-fitting functions in these algorithms.

Hygiena's foodproof® assays utilize unsupervised machine-learning algorithms. Hygiena currently favors this technique for future assay algorithm development, where there is a small amount of data available for training.

Experimental Algorithms

Effective in BAX System Q7 Software algorithm version 4.22.0.9784 and beyond, Hygiena will introduce new features to optionally incorporate our experimental algorithms and account for computational instabilities in different computer processors. Certified algorithms are unchanged. Users may choose to select between Certified or Experimental algorithms. These new features are currently limited to BAX System Real-Time *Salmonella* and Real-Time *E. coli* EXACT Assays. Customers may choose to have a messaging feature to aid them in better understanding why a call may have been made. For example, in a PCR curve with a slight positive shape, but the algorithm interprets it as negative, a user may see messages such as "Weak signal. Signal-to-noise ratio levels under minimum allowable threshold."

Customers may also choose to enable experimental algorithms for foodproof assays. If certified algorithms are selected, only a question mark will be displayed on the well.

Frequently Asked Questions:

- **Will I see a difference in my Ct value when deploying different types of algorithms?**
These updates do not affect a graph or curve; however, we have found changes in the Ct values when moving from the older models to the new machine-learning models. This is because we use different functions to calculate the Ct values. The old functions were slow and computationally problematic when running large amounts of data and had trouble tracing difficult curve types.

  Hygiena performed extensive research on the changes in Ct values when upgrading our models. We never found a consistent shift in Ct values in any direction (increased Ct or decreased Ct). We also found that these differences in Ct values were statistically negligible

when compared to the older models. They do not significantly impact results that would change any validations or claims.

- **Are these new algorithms less accurate?**
  When we changed the curve-fitting functions to estimate, after extensive validations, we found that our new models always performed better than the old statistical models and more accurately identified the correct Ct values. In addition to the Ct value, our enhanced statistics allow for improved confidence in both weak and strong positive results.