applied biosystems

Axiom™ Microbiome Solution

USER GUIDE

Catalog Numbers 902910, 902904, and 902903

Publication Number 703408

Revision 2

ThermoFisher SCIENTIFIC

For Research Use Only. Not for use in diagnostic procedures.

Manufacturer:
Thermo Fisher Scientific Baltics UAB
V.A. Graiciuno 8, LT-02241
Vilnius, Lithuania

Products: Axiom™ Microbiome Reagent Kit

Manufacturer:
Affymetrix Pte Ltd
7 Gul Circle #2M-01
Keppel Logistics Building
Singapore 629563

Products: Axiom™ Microbiome Array Plates

The information in this guide is subject to change without notice.

Revision history

Table 1 Revision history of Pub. no. 703408

Revision	Date	Description
2	04 September 2018	Baseline for revision history. Updated to the current document template, with associated updates to trademarks, logos, licensing, and warranty. Updated to reflect that Axiom™ Reference gDNA 103 has been removed from the reagent kit and has been made available for purchase separately.

Corporate entity

Life Technologies | Carlsbad, CA 92008 USA | Toll Free in USA 1 800 955 6288

Contents

Section 1:
Axiom™ Microbiome Assay Protocol

1 Introduction

About the Axiom™ Microbiome Assay

Axiom™ Microbiome Array is a microarray developed to identify microorganisms present in a sample and was designed in collaboration with Lawrence Livermore National Laboratory (LLNL). In contrast to genotyping and gene expression microarrays, which interrogate SNPs or genes with a defined set of probes, all of the probes on Axiom Microbiome Array are used by the analysis algorithm ("Axiom™ MiDAS algorithm stages" on page 67) to identify the microorganisms present in a sample.

The Axiom Microbiome Array uses the same Axiom™ 2.0 Assay as all other Axiom products. Variations or deviations specific to Axiom Microbiome Array, including the use of a cDNA input as an entry point into the assay, assay-specific labware, reagents, and control samples are detailed in this guide.

Axiom Microbiome Array contains probes for approximately 12,000 microbial species, enabling for high coverage, cost-effective studies. It leverages the Axiom 2.0 Assay for target preparation and array processing:

- For 96-format arrays, target preparation can be done manually or with the use of automation using Applied Biosystems™ NIMBUS™ Target Preparation Instrument. The NIMBUS Instrument contains methods for DNA amplification, fragmentation, purification, and resuspension of the target in hybridization cocktail.
- For 24-array format, target preparation can be done manually.
- Hands-free processing of all array plate formats are performed on the GeneTitan™ MC Instrument.

Prerequisites

As Axiom Microbiome Array leverages the Axiom 2.0 Assay for target preparation, users should review and become familiar with the appropriate documentation for the variation of the workflow(s) that the user intends to use (see "Related documentation" on page 91). Analysis of Axiom Microbiome Array requires the use of Axiom™ Microbial Detection Analysis Software (MiDAS), a set of analysis tools specific to this product, rather than Axiom™ Analysis Suite used for Axiom genotyping products.

2 DNA preparation and requirements

The general requirements for genomic DNA (gDNA) sources and extraction methods are described in this chapter. The success of this assay requires uniform amplification of the genome starting with relatively intact gDNA. To achieve this, the gDNA must be of high quality, and must be free of contaminants that may affect the enzymatic reactions to be performed.

For Axiom™ Microbiome Array, use Axiom™ 2.0 Reagent Kit (Cat. No. 901758) or Axiom™ Microbiome Reagent Kit 4x24 Reactions (Cat. No. 902910). Axiom™ Reference Genomic DNA 103 (Cat. No. 951957) is available for purchase separately. The Axiom Microbiome Reagent Kit 4x24 Reactions includes 3 tubes of the Genomic DNA control such that 1 control sample can be used for each 24 sample target preparation and array hybridization. This human DNA control meets the requirements outlined below. The size and purity of sample gDNA can be compared with those of the control DNA to assess sample quality. This Genomic DNA control ***must*** be used as an experimental positive control throughout target preparation and must be included on each Axiom Microbiome Array Plate as it provides assurance that both assay and array processing steps are working properly. Information on the performance of the plate positive control is necessary to investigate or to provide warranty replacement for any product performance complaint.

Assay performance may vary for gDNA samples that do not meet the general requirements described below. However, the reliability of any given result should be assessed in the context of overall experimental design and goals.

The genomic DNA requirements and preparation are described in the following sections:

Axiom™ Microbiome Array can be used to analyze viral RNA genomes composed of ssRNA or dsRNA. To interrogate RNA samples, cDNA from RNA samples must be generated and used as an input into the recommended Axiom™ target preparation workflows. To generate cDNA, RNA can be reverse transcribed to yield input amenable to Axiom target preparation using the protocol outlined in Chapter 3, "cDNA synthesis for RNA samples" on page 21.

Note: The generation of cDNA from RNA samples is only recommended for Axiom Microbiome Array target generation for the interrogation of RNA viruses on this platform. This is *not* a supported sample preparation for Axiom genotyping arrays.

Sources of genomic DNA

Genomic DNA extracted from stool samples has been successfully tested in laboratories at Thermo Fisher Scientific and meets the general requirements as outlined in this user guide. Success with other types of samples will depend on quality (degree of degradation, level of purity, etc.) and quantity of gDNA extracted.

Note: DNA derived from formalin-fixed paraffin-embedded (FFPE) blocks should not be used with this assay.

Table 1.1 gDNA sample types

Sample type	Volume per well	Input mass per well	gDNA concentration
Stool gDNA	20 µL	50 ng	2.5 ng/µL

General requirements

- Starting DNA must be double-stranded for the purpose of accurate concentration determination.
- DNA must be of high purity.
  DNA should be free of DNA polymerase inhibitors. Examples of inhibitors include high concentrations of heme (from blood) and high concentrations of chelating agents (i.e., EDTA). The gDNA extraction/purification method should render DNA that is generally salt-free because high concentrations of particular salts can also inhibit enzyme reactions. DNA purity is indicated by OD_{260}/OD_{280} and OD_{260}/OD_{230} ratios. The OD_{260}/OD_{280} ratio should be between 1.8 and 2.0 and the OD_{260}/OD_{230} ratio should be greater than 1.5. We recommend that DNA samples that do not meet these criteria be cleaned up as described under "Genomic DNA cleanup" on page 13.
- DNA must not be degraded.
  The approximate average size of gDNA may be assessed on a 1% agarose gel using an appropriate size standard control. Approximately 90% of the DNA must be greater than 10 Kb in size. Control DNA can be run on the same gel for side-by-side comparison.

Special requirements

Preamplification area

Precautions are required when manipulating genomic DNA to avoid contamination with foreign DNA amplified in other reactions and procedures. It is recommended that genomic DNA manipulations are performed in a dedicated preamplification room or area separate from the main laboratory.

This pre-amplification area should have a dedicated set of pipettes and plasticware. If no dedicated area is available, use of a dedicated bench or a dedicated biosafety hood and dedicated pipettes is suggested. If no dedicated bench or biosafety hood is available, a set of dedicated pipettes is recommended.

Ideally, this preamplification area would be separate from the amplification staging area described below.

Amplification staging area

Precautions are required when setting up amplification reactions to avoid contamination with foreign DNA amplified in other reactions and procedures. It is recommended that amplification reaction set up is performed in a dedicated amplification staging area separate from the main laboratory.

This amplification staging area should have a dedicated set of pipettes and plasticware. If no dedicated amplification staging area is available, use of a dedicated bench or a dedicated biosafety hood and dedicated pipettes is suggested. If no dedicated bench or biosafety hood is available, a set of dedicated pipettes is recommended.

Fume hood

At certain steps in the protocol we recommend the use of adequate local or general ventilation to keep airborne concentrations low.

A fume hood is suggested as a way to achieve the desired concentration. Thus, a fume hood is strongly recommended for several steps of this assay.

Labware consumables specific to the Axiom™ Microbiome Assay

For Axiom Microbiome applications, only use the MicroAmp™ Clear Adhesive Film (Cat. No. 4306311). The use of the Bio-Rad Microseal 'B' Film seal, which is also an approved seal for Axiom Genotyping applications, has not been validated for Axiom Microbiome use.

For Axiom Microbiome manual workflows, the Eppendorf 96 Deep-well Plate, 2,000 µL is used in lieu of the ABgene 96 Square Well Storage Plate, 2.2 mL used in Axiom genotyping workflows.

Assessing the quality of genomic DNA using 1% agarose E-gels

We recommend this quality control step to asses the quality of the gDNA prior to starting the assay.

Equipment and reagents recommended

Table 1.2 E-Gel® and reagents required

Item	Source
Mother E-Base Device	EB-M03
Daughter E-Base Device	EB-D03
E-Gel® *48* 1% agarose gels	G8008-01
*Redi*Load™	750026
E-Gel® 96 High Range DNA Marker	12352-019

Guidelines for preparing the Genomic DNA Plate for gel analysis

- Loading a DNA mass of 10 ng to 20 ng per well is recommended. If lower amounts are loaded, omission of the loading dye is recommended in order to improve visualization. Loading ≥25 ng gDNA per well can improve the image.
- Add 3 µL of 0.1X of *Redi*Load dye to each sample.
- Bring each sample to a total volume of 20 µL using H_2O (for example, if the volume of genomic DNA is 5 µL, add 3 µL of *Redi*Load, and bring to 20 µL total by adding 12 µL of H_2O).
- Seal, vortex, and spin.

To run a 48-lane 1% Agarose E-Gel:

1. Power on for E-Base (red light).
2. Push the Power/Prg button to make sure the program is at EG mode (not EP).
3. Adjust the run time to ~27 min.
4. Insert the 48 well 1% Agarose E-Gels into the slot.
5. Remove the combs.
6. Load 20 µL from the above plate onto two 48 well 1% agarose E-Gels.
7. Load 15 µL of diluted High Range DNA Marker (1:3 dilution or ~0.34X from stock) into all marker wells (as needed).
8. Fill all empty wells with water.
9. Push the Power/Prg button again (it will change from red to green).

When run time is reached (the ladder band reaches the end of the lane), the system will automatically shut off. The gel is then ready for imaging.

Figure 1.1 shows gel images of intact gDNA (that is suitable for use in the Axiom 2.0 Assay) and degraded gDNA samples. Customers whose gDNA is degraded (similar to the image in Figure 1.1) should perform a test experiment to investigate the performance of their samples in the Axiom 2.0 Assay prior to beginning any large scale projects.

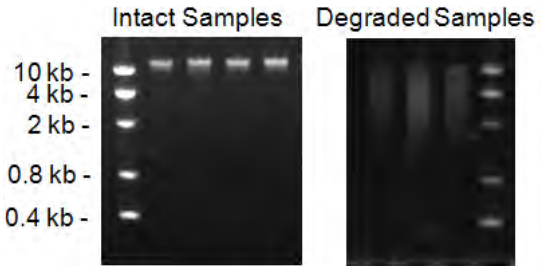


Figure 1.1 Gel images showing intact gDNA and degraded gDNA

Genomic DNA extraction/purification methods

Genomic DNA extraction and purification methods that meet the general requirements outlined above should yield successful results. Methods that include boiling or strong denaturants are not acceptable because the DNA would be rendered single-stranded and can no longer be accurately quantitated using a PicoGreen-based assay.

Genomic DNA cleanup

If a gDNA preparation is suspected to contain inhibitors, the following cleanup procedure can be used:

1. Add 0.5 volumes of 7.5 M NH_4OAc, 2.5 volumes of absolute ethanol (stored at –20°C), to gDNA.
2. Vortex and incubate at –20°C for 1 hour.
3. Centrifuge at 12,000 x *g* in a microcentrifuge at room temperature for 20 minutes.
4. Remove supernatant and wash pellet with 80% ethanol.
5. Centrifuge at 12,000 x *g* at room temperature for 5 minutes.
6. Remove the 80% ethanol and repeat the 80% ethanol wash one more time.
7. Resuspend the pellet in reduced EDTA TE Buffer (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA).

Quantitate the genomic DNA

1. Gently vortex (50% maximum) and spin the gDNA.
2. Recommendation: quantitate each sample (e.g., using the Quant-iT™ Picogreen® dsDNA Kit).
3. At this point, you can:
   - Prepare cDNA: proceed to Chapter 3, "cDNA synthesis for RNA samples" on page 21, or
   - Proceed to "DNA sample preparation" on page 14.

DNA sample preparation

This step needs to be performed prior to proceeding with the DNA amplification stages. The genomic DNA (gDNA) you will process using the Axiom 2.0 Assay should meet the general requirements listed earlier in this chapter. For Axiom Microbiome a total of 50 ng of gDNA is required for DNA samples. If using Axiom Microbiome Array for the interrogation of RNA genomes, samples should be prepared as written in Chapter 3, "cDNA synthesis for RNA samples" on page 21.

To prepare gDNA:

"1: Thaw samples and control" on page 15

"2: Dilute the DNA Samples" on page 16

"3: Aliquot the diluted samples, Reference Genomic DNA 103 Control, and NTC" on page 16

"4: Freeze or proceed" on page 18

"5: Create a GeneTitan Array Plate Registration file" on page 20

Equipment, consumables, and reagents required

The equipment and consumables listed in Table 1.3 are required for gDNA Preparation.

Table 1.3 Equipment and reagents required for genomic DNA preparation

Quantity	Item
As required	Adhesive Film • MicroAmp™ Clear Adhesive Film (Cat. No. 4306311)
1	Ice Bucket, filled with ice
1 each	Pipettes: • Single Channel P10 or P20 • Optional: Multichannel P10 or P20
As required	Pipette tips
1	Deep-well Plate[1] If running 24F or 96F manual workflow: – Eppendorf 96 Deep-well Plate, 2,000 µL If running Axiom NIMBUS automated workflow: – Round Deep-well Plate (Axygen)
1	Plate centrifuge
1	Plate spectrophotometer
1	Vortexer

[1] Refer to Table 1.5 on page 17 for plate part numbers.

Reagents

The reagents listed in Table 1.4 are required for DNA sample preparation.

Table 1.4 Reagents required for genomic DNA preparation

Reagent
Axiom Reference Genomic DNA 103, used as a positive control (Cat. No. 951957), –20°C
User-supplied
Reduced EDTA TE Buffer (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA) (Cat. No. 75793)
Buffer control used for elution of gDNA samples

Note: The Axiom Microbiome Reagent Kit 4 x24 Reactions (Cat. No. 902910) contains 3 vials of Axiom Reference Genomic DNA 103 (Part No. 901012).

1: Thaw samples and control

Thaw the components below to room temperature:

- gDNA samples.
- cDNA samples. See Chapter 3, "cDNA synthesis for RNA samples" on page 21 for instructions on how to prepare cDNA from viral RNA samples.
- gDNA positive control samples. For Axiom Microbiome, Axiom Reference Genomic DNA 103 (from Axiom™ Microbiome Reagent Kit 4x24 Reactions or Cat. No. 951957) must be used for each array plate to provide both Assay QC and Array QC metrics.
- No Template control (NTC). It is necessary to run one no template control (NTC) reaction. It is recommended that the user run the same buffer used for elution during their gDNA extraction. Alternatively, Reduced EDTA TE Buffer (Cat. No. 75793) can be run as a negative control.

To Thaw, either:

- Place items on benchtop for one hour

  Or

- Thaw in a water bath:

1. Fill a small plastic dish with Millipore water. Do not immerse the sample plate or tube when placing it in the bath.
2. Thaw the sealed sample plate and/or sample tube for a half-hour.
3. Remove the sample plate and/or sample tube from the water bath and wipe-dry using lab wipes. Ensure the outside is completely dry before opening the sample plate or tube to minimize any contamination, which can lead to reaction failure.

2: Dilute the DNA Samples

1. Gently vortex (50% maximum) and spin the DNA Samples and Axiom Reference Genomic DNA 103 (positive control sample).
2. Using Reduced EDTA TE Buffer, dilute each gDNA sample to a concentration of 2.5 ng/µL.
3. When including cDNA templates (see Chapter 3, "cDNA synthesis for RNA samples" on page 21) first transfer 2.5 µL of Reduced EDTA TE Buffer to the sample plate, then add 17.5 µL of cDNA template generated.
4. Seal, vortex, and spin.

Note: Do *not* dilute the Axiom Reference Genomic DNA 103 control. It is already at a working concentration.

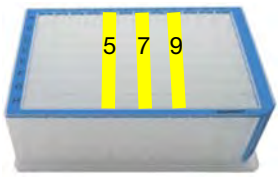
3: Aliquot the diluted samples, Reference Genomic DNA 103 Control, and NTC

Next, the samples and controls are transferred to a deep-well plate for target preparation:

- For 24-Format manual target preparation, use the Eppendorf 96 Deep-well Plate, 2,000 µL Eppendorf Cat. No. 951033481. Refer to "Aliquot diluted samples, controls, and Reference Genomic DNA 103 to columns 5, 7, and 9 of the Eppendorf Deep-well Plate for 24-array format manual workflow" below for details on how to plate your starting material.
- For the 96-Format manual target preparation, use the Eppendorf 96 Deep-well Plate, 2,000 µL Eppendorf Cat. No. 951033481.
- For the 96-Format automated target preparation on the Applied Biosystems NIMBUS Instrument, use the Round Deep-well Plate. Axygen Cat. No. P-DW-20-CS-S.

IMPORTANT! The manual and automated target preparation options require different 96- well deep-well plates for gDNA or cDNA plating.

Table 1.5 Deep-well plate information for Axiom™ Microbiome Assays

Axiom™ Assay	Format	Deep-well plate	Plate name	Manufacturer Information
24-array format manual workflow	4x24		Eppendorf 96 Deep-well Plate, 2,000 µL	Eppendorf Cat. No. 951033481
96-array format manual workflow	96		Eppendorf 96 Deep-well Plate, 2,000 µL	Eppendorf Cat. No. 951033481
96-array format automated workflow–NIMBUS	96		Round Deep-well Plate Part of the Axiom™ 96 Consumables Kit for Applied Biosystems® NIMBUS® 2.0 (Kit Cat. No. 902907, Plate Part No. 203028).	Axygen® Cat. No. P-DW-20-C-S

Aliquot diluted samples, controls, and Reference Genomic DNA 103 to columns 5, 7, and 9 of the Eppendorf Deep-well Plate for 24-array format manual workflow

1. Aliquot 20 µL of each diluted gDNA, and/or no template control samples to columns 5, 7, and 9. This should be the equivalent of 50 ng of gDNA, as required.
2. Aliquot 20 µL of the Reference Genomic DNA 103 control to the plate into well H09.
3. When including cDNA templates (see Chapter 3, "cDNA synthesis for RNA samples" on page 21) first transfer 2.5 µL of Reduced EDTA TE Buffer to the sample plate. Then add 17.5 µL of cDNA template generated.
4. Seal and spin.

Note: Thermo Fisher Scientific requires including Reference Genomic DNA 103 as a positive control and the use of a no template control on each plate

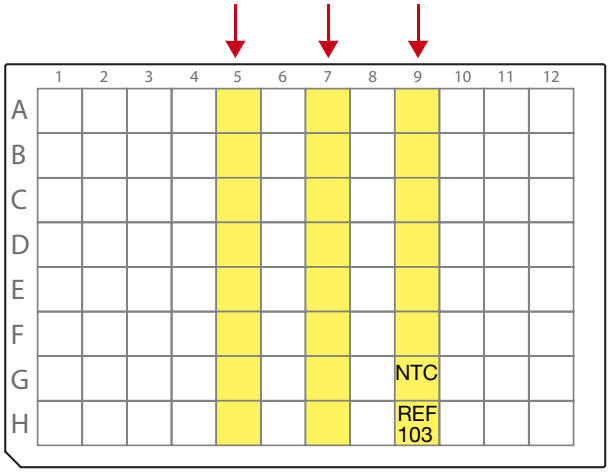


Figure 1.2 Aliquot diluted gDNA samples to columns 5, 7, and 9 only

4: Freeze or proceed

At this point you can:

- Store the sample plate at –20°C, or
- Proceed to DNA Amplification to begin target preparation. Reference the appropriate user guide as outlined below on Table 1.6.

Note: You can leave the DNA sample plate at room temperature if proceeding immediately to DNA Amplification.

24-array format manual workflow

To continue with 24-array format manual target preparation, refer to *Axiom™ 2.0 Assay 24-Array Format Manual Workflow User Guide*, Pub. No. 703335, Chapter 3. Ensure that your gDNA and/or cDNA samples are plated on the Eppendorf 96 Deepwell Plate, 2,000 µL. Please refer to Table 1.5 on page 17 for details on how to obtain this plate.

All steps will be performed as written except the use of the different 96 square deep-well plate noted above. Please disregard sample input requirements for other array formats and follow the input requirements as outlined in this section.

96-array format manual workflow

To continue to with 96-format manual target preparation refer to *Axiom™ 2.0 Assay Manual Workflow User Guide*, Pub. No. 702990, Chapter 3. Ensure that your gDNA and/or cDNA samples are plated on the Eppendorf 96 Deep-well Plate, 2,000 µL. Please refer to Table 1.5 on page 17 for details on how to obtain this plate.

All steps will be performed as written except the use of the different 96 Square Deep-well plate noted above. Please disregard sample input requirements for other array formats and follow the input requirements as outlined in this section.

96-array format automated workflow on Applied Biosystems™ NIMBUS™

To continue to with 96-format target preparation on Applied Biosystems™ NIMBUS™ refer to *Axiom™ 2.0 Assay 96-Array Format Automated Workflow User Guide—Applied Biosystems™ NIMBUS™*, Pub. No. 703349, Chapter 3. Ensure that your gDNA and/or cDNA samples are plated on the round deep-well plate (Axygen). See Table 1.5 on page 17 for details on how to obtain this plate.

All steps will be performed as written. Please disregard sample input requirements for other array formats and follow the input requirements as outlined in this section.

IMPORTANT! For Adhesive Film to seal plates, only use the recommended Plate Seal - MicroAmp™ Clear Adhesive Film: Thermo Fisher Scientific Cat. No. 4306311. Other seals have not been validated for use with this workflow.

Table 1.6 Appropriate Axiom™ 2.0 Assay user guide for Axiom™ Microbiome assays

Axiom assay	User guide	Chapter
24-array format manual workflow	*Axiom™ 2.0 Assay 24-Array Format Manual Workflow User Guide*, Pub. No. 703335	Chapter 3: Axiom™ 2.0 Assay for 24 Samples: Preparation Before You Start Chapter 4: Axiom™ 2.0 Assay 24 Format Manual Target Preparation
96-array format manual workflow	*Axiom™ 2.0 Assay Manual Workflow User Guide*, Pub. No. 702990	Chapter 3: Axiom™ 2.0 Assay: Preparation Before You Start Chapter 4: Axiom™ 2.0 Assay: Manual Target Preparation
96-array format automated workflow on the Applied Biosystems™ NIMBUS™ Instrument	*Axiom™ 2.0 Assay 96-Array Format Automated Workflow User Guide—Applied Biosystems™ NIMBUS™*, Pub. No. 703349	Chapter 3: Target Preparation with the Applied Biosystems™ NIMBUS™ Target Preparation Instrument

5: Create a GeneTitan Array Plate Registration file

IMPORTANT! It is very important to create and upload a GeneTitan Array Plate Registration file with your sample information prior to loading the array plate and hyb tray in the GeneTitan Instrument. We recommend that you create (but not upload) this file at the same time you prepare your plate of gDNA and/or cDNA. When your samples are ready for hybridization, you will scan the array plate barcode and upload the file to GeneChip Command Console (GCC).

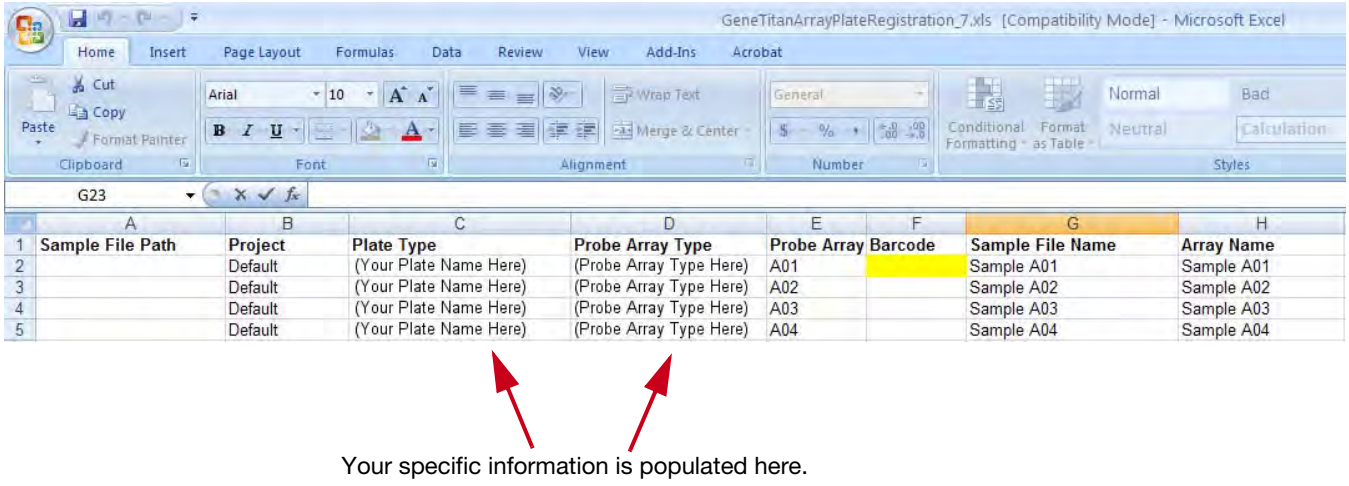
GeneTitan Array Plate Registration files contain information that is critical for:

- Data file generation during imaging.
- Tracking the experimental results for each sample loaded onto an array plate.

See Figure 1.3 for a screen shot showing an example of a batch registration file.

1. Open AGCC Portal → Samples, and select:
   a. GeneTitan Array Plate Registration.
   b. The array plate format.
   c. Click **Download**.
2. Enter a unique name for each sample and any additional information.
3. Save the file.

The array plate barcode will not be scanned until you are ready to load the array plate and samples onto the GeneTitan MC Instrument for processing.



	A	B	C	D	E	F	G	H
1	**Sample File Path**	**Project**	**Plate Type**	**Probe Array Type**	**Probe Array**	**Barcode**	**Sample File Name**	**Array Name**
2		Default	(Your Plate Name Here)	(Probe Array Type Here)	A01		Sample A01	Sample A01
3		Default	(Your Plate Name Here)	(Probe Array Type Here)	A02		Sample A02	Sample A02
4		Default	(Your Plate Name Here)	(Probe Array Type Here)	A03		Sample A03	Sample A03
5		Default	(Your Plate Name Here)	(Probe Array Type Here)	A04		Sample A04	Sample A04

Figure 1.3 Example of a GeneTitan Array Plate Registration file

3 cDNA synthesis for RNA samples

Axiom™ Microbiome can be used to analyze viral RNA genomes composed of ssRNA or dsRNA. To interrogate RNA samples, cDNA from RNA samples must be generated and used as an input into the recommended Axiom™ target preparation workflows. To generate cDNA, RNA can be reverse transcribed to yield input amenable to Axiom target preparation using the protocol outlined below. The protocol outlined below uses the SuperScript® VILO™ cDNA Synthesis Kit (Cat. No. 11754050). For general guidelines on the use of this kit, please refer to the Product Information Sheet: SuperScript VILO cDNA Synthesis Kit.

Note: The generation of cDNA from RNA samples is only recommended for Axiom™ Microbiome Array target generation for the interrogation of RNA viruses on this platform. The use of cDNA with Axiom™ genotyping arrays is not validated

General RNA requirements

The quality of the RNA input is essential to the overall success of the analysis. Since the most appropriate protocol for the isolation of RNA can be source dependent, we recommend using a protocol that has been established for the source used.

When using a commercial kit, follow the manufacturer's recommendations for RNA isolation.

RNA sample purity and concentration

- The absorbance should be checked at 260 and 280 nm for determination of sample purity.
- The A_{260}/A_{280} ratio should be close to 2.0 for pure RNA.
- Recommendation: Quantitate each RNA sample (e.g., using Quant-iT™ RiboGreen® RNA Assay Kit)

Equipment, consumables, and reagents required

Equipment and consumables

Table 1.7 Consumables required for cDNA synthesis

Consumable	Supplier	Cat. No.
Pipettes—P2, P10, P20	Various	—
Filtered pipette tips	Various	—

Reagents

Table 1.8 Reagents required for cDNA synthesis

Reagent	Supplier	Cat. No.
SuperScript™ VILO™ cDNA Synthesis Kit (50–20 µL reactions)	Thermo Fisher Scientific	11754-050
Reduced EDTA TE Buffer (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA)		75793
DEPC-treated Water	Various	—

General guidelines

- Disposable lab coats, gloves and shoe coverings should be worn at all times while in a clean room environment. Discard after use.
- Ensure workspace and materials are free of RNases:
  - RNaseZap™ (Cat. No. 9780) can be used to decontaminate pipettes, benchtop and plastic surfaces.
  - Ensure all labware (pipette tips, tubes, etc.) are certified RNase-free.
  - If possible, designate an 'RNA only' (RNase free) workspace including designated bench space, pipettes, pipette tips, reagents and consumables.
  - Gloves should be worn at all times and changed often. Once gloves come into contact with human skin or hair, pipettes, refrigerator doors, door knobs, lab benches, etc. they are no longer considered RNase free.
- Only filtered pipette tips should be used to help prevent pipette contamination.
- RNA samples should be stored and thawed properly:
  - Store at –80°C when not in use.
  - Thaw in an ice bucket and keep on ice throughout cDNA synthesis reaction preparation. Return samples immediately to –80°C freezer after use.
  - Minimize freeze/thaw cycles to help prevent RNA degradation. Aliquoting smaller volumes of an RNA sample into multiple tubes helps to minimize the number freeze/thaw cycles per tube.

RNA sample requirements

Recommended RNA positive control

To ensure the generation of cDNA from RNA a positive control reaction is recommended. 1 ng of RNA of Mumps virus RNA, strain Enders (ATCC® VR--106D™) can be used as a control for the generation of a cDNA input for generation of target to be interrogated by Axiom Microbiome Arrays.

Input requirements for 20 µL reaction

- Minimum total RNA input: 1 ng
- Maximum total RNA input: 2.5 µg
- Maximum RNA sample volume: 14 µL

1: Reagent and sample preparation

Prepare reagents and samples in a clean room environment:

1. 5X VILO™ Reaction Mix—thaw on ice.
   - Vortex and spin before use.
2. RNA sample—thaw on ice.
   - Vortex and spin before use.
3. 10X SuperScript™ Enzyme Mix—leave in freezer, do not thaw.
   - Flick tube 3 times and spin immediately before use.

2: Reaction preparation

Prepare reaction in a clean room environment:

1. Combine the following components in a tube on ice for a single 20 µL reaction. If processing multiple 20 µL reactions, prepare a master mix without RNA:

Table 1.9 Reverse Transcription Master Mix

Component	Volume
5X VILO™ Reaction Mix	4 µL
10X SuperScript™ Enzyme Mix	2 µL
RNA Sample (up to 2.5 µg)	x µL
DEPC-treated water	14 - x µL

   a. Using a P20 pipette, add (14 – x) µL of DEPC-treated water.
   b. Using a P10 pipette, add 4 µL of 5X VILO Reaction Mix.
      - Vortex and spin before use.
   c. Using a P2 pipette, add 2 µL 10X SuperScript Enzyme Mix.
      - Flick tube 3 times and spin before use.
   d. Using a P10 or P20 pipette, add x µL of RNA Sample (up to 2.5 µg total RNA).
      - Vortex and spin before use.
2. Gently mix tube and spin.

3: Incubation

Move prepared reaction out of the clean room on ice.

1. Incubate at 25°C for 10 minutes.
2. Incubate 42°C for 120 minutes.
3. Incubate 85°C for 5 minutes.
4. Freeze or Proceed:
   At this point you can:
   - Store the cDNA sample at –20°C, or
   - Proceed to sample plate preparation as detailed in Chapter 2, *Step* "3: Aliquot the diluted samples, Reference Genomic DNA 103 Control, and NTC" on page 16.

IMPORTANT! For interrogation of cDNA templates 17.5 µL of the Reverse transcription reaction will be added to 2.5 µL Reduced TE Buffer that has been added to the sample plate. The remaining 2.5 µL of this reaction can be saved and used for other applications.

IMPORTANT! If processing an Axiom plate with both gDNA and cDNA samples, add the gDNA samples in a clean room first then transfer the plate into a non-clean room environment to add the cDNA samples.

Section 2:
Axiom™ Microbial Detection Analysis Software (MiDAS)

1 Introduction

Overview

Axiom™ Microbial Detection Analysis Software (MiDAS) is a software package designed for analysis of Axiom™ Microbiome Array data using the Composite Likelihood Maximization (CLiMax) algorithm. Axiom MiDAS output is a list of microbial organisms likely to be present in a sample. The software provides tools to compare families and species detected in different samples and reports Axiom MiDAS results for individual samples in both table and graphical formats.

Highlights

- View results data that summarize microbial content in table and graphical formats.
- Generate a list of microbial target sequences that most likely comprise the sample.
- Directly access external NCBI databases.

Software and hardware requirements

Table 2.1 Axiom™ MiDAS hardware and software specification

64-bit operating system	Speed	Memory	Available disk space[1]
Microsoft Windows® 7 (64 bit) Professional with Service Pack 1	3 GHz Intel Pentium Quad Core Processor	16 GB RAM	30 GB HD + data storage (~50 GB/plate)
Microsoft Windows® 10 (64 bit) Professional	3 GHz Intel Pentium Quad Core Processor	16 GB RAM	30 GB HD + data storage (~50 GB/plate)

[1] Minimum storage requirements are for a single run. Total storage space should include additional space for data storage of input and output files from current and previously completed analyses. In addition, you must have a minimum of 5 GB of free space on your C: drive to run an analysis. A batch name folder is auto-generated during the analysis process. This folder includes all the necessary files needed to view your analysis results in the Viewer.

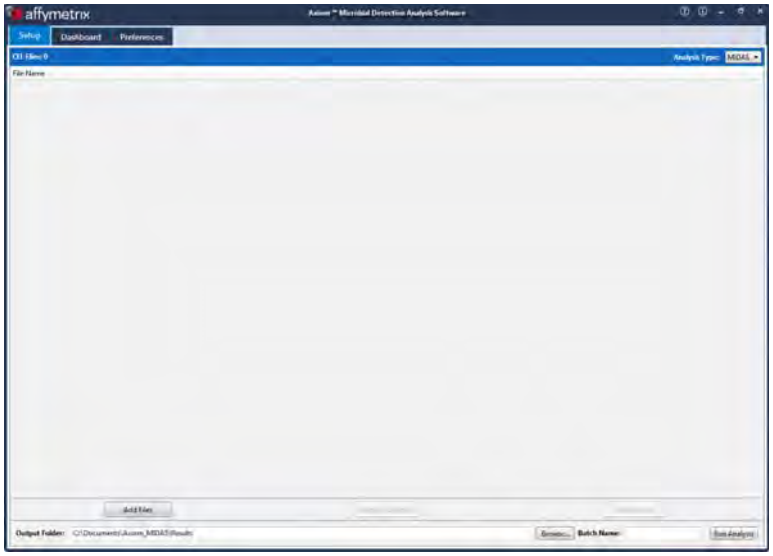
Note: To maximize the speed of your analysis run, make sure you close all other open and/or running programs before running your analysis.

Installation instructions

1. Go to thermofisher.com, locate and download the zipped Axiom MiDAS software package.
2. Unzip the file, then double-click **Axiom_MiDAS_Setup.exe**.
3. Follow the on-screen instructions to complete the installation.

Starting Axiom™ MiDAS

1. Click **Start** → **All Programs** → **Affymetrix** → **Axiom MiDAS**.
   The following window appears:



Preferences

Click the **Preferences** tab to view assigned path locations, edit proxy settings, and check or auto-check for library updates (Figure 2.1).

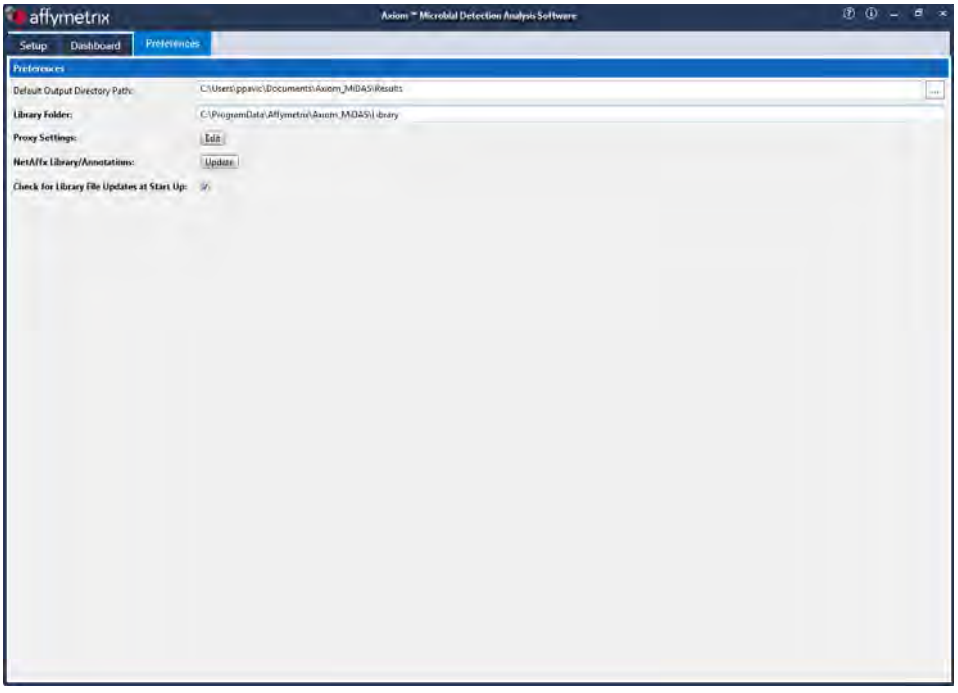


Figure 2.1 Main Preferences window

Using preferences

Default output directory path

To change the default output directory path:

1. Click ... .

   A **Select Default Output Folder** window appears.

2. Navigate to a new folder location as you normally would, then click **Select Folder**.

   The output directory path is now changed.

Library folder

This path is set during the Axiom MiDAS installation and cannot be changed.

Proxy Settings

If your system has to pass through a proxy server before it can access the NetAffx™ server (internet), click Edit .

The following window appears (Figure 2.2):

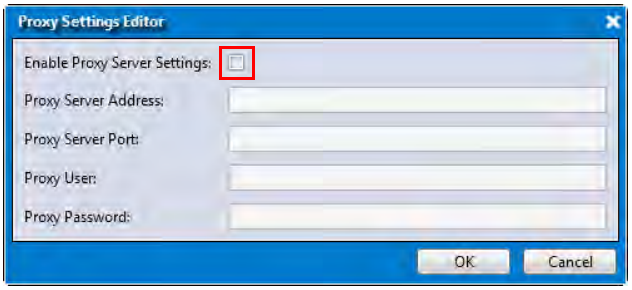


Figure 2.2 Proxy Settings Editor window

3. Click the **Enable Proxy Server Settings** check box (Figure 2.2), then contact your IT department for help with completing the required text fields.
4. Click **OK**.

Updating NetAffx™ Library/Annotations

1. Click Update .

   The NetAffx Login window appears (Figure 2.3).

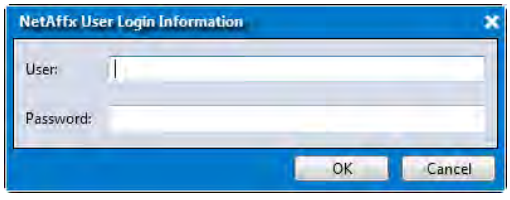


Figure 2.3 NetAffx Login window

2. Enter your **User** name and **Password**, then click **OK**.

   Note: If you are unable to connect to the NetAffx Download Center, make sure you have entered the correct NetAffx User name and Password, have an active internet connection, and proper proxy server settings.

   - If you do not have a NetAffx account, go to www.affymetrix.com, click **NetAffx**, then click **Register**.
   - After registering, try to login to the NetAffx Download Center again.

   The NetAffx Update window appears (Figure 2.4).

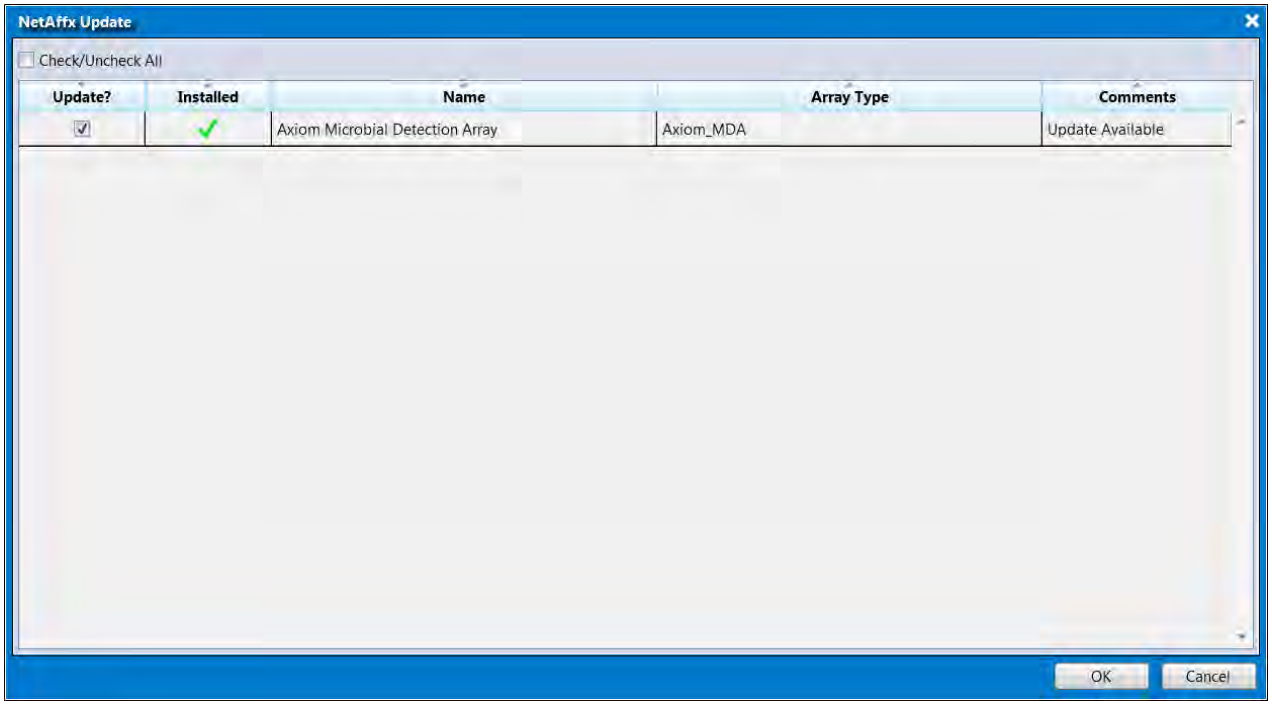


Figure 2.4 NetAffx Update window

3. You must select the check box(es) that correspond with the type of CEL files you want to analyze.

   Table 2.2 Click the **Check/Uncheck All** check box to select/deselect all the listed check boxes.

4. Click **OK**.

   An Installing Updates progress bar appears. As each applicable update completes its download, a green check mark appears inside its **Installed** column.

Check for Library File Updates at Start Up

1. Click the **Check for Library File Updates at Start Up** check box (Figure 2.5) to enable automatic library file update alerts each time you launch the Axiom MiDAS application (recommended).

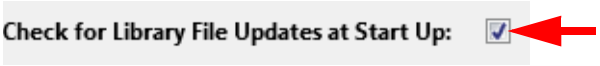


Figure 2.5 Auto-update notifications check box

Downloading library files manually

Use this downloading option if you have successfully installed the Axiom MiDAS software, but cannot download the library files through the software.

1. Go to: www.thermofisher.com
2. In the search bar at the top center of the page, enter *Axiom Microbiome*.
3. From the results, select the appropriate Microbiome array format.
4. On the Axiom Microbiome product page, scroll down to the **Support files** section.
5. Click the Axiom Microbiome Library File link.

   The Customer Login window appears.
6. Login as you normally would, then click **Sign In**.

   An Internet Explorer dialog window appears.
7. Click **Save As**.
8. Note the save (folder) location.
9. Leave the .7z file name as is, then click **Save**.

Unzipping the Library Files

IMPORTANT! The zipped Axiom MiDAS library file can only be unzipped using *7-zip*. This file archiving tool can be downloaded for free at: www.7-zip.org

1. If 7-zip is not currently installed on your system, go to: www.7-zip.org
2. Click on the download link that matches your operating system.
3. **Run** or **Save** the executable file.
4. After 7-zip is installed on your system, double-click on its icon to open it.

   The following window appears (Figure 2.6).

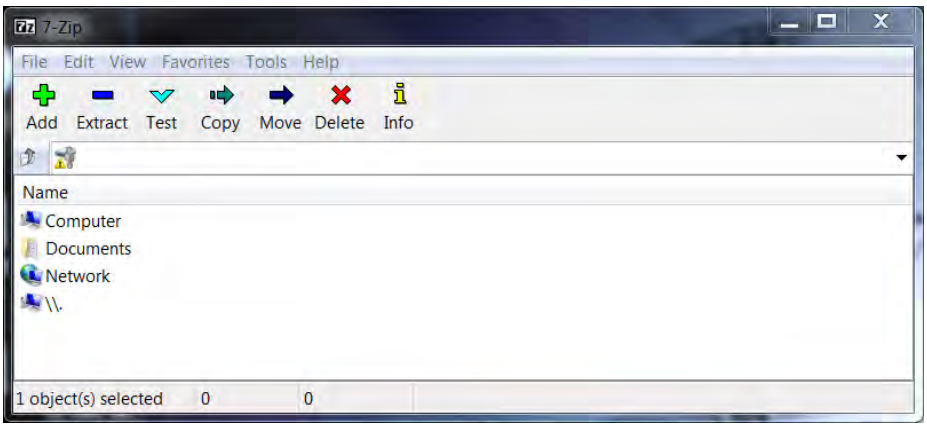


Figure 2.6 7-zip main window

5. Use 7-zip's main window to locate the Axiom MiDAS library zip file.

6. Click to highlight it, then click **Extract** (Figure 2.7).

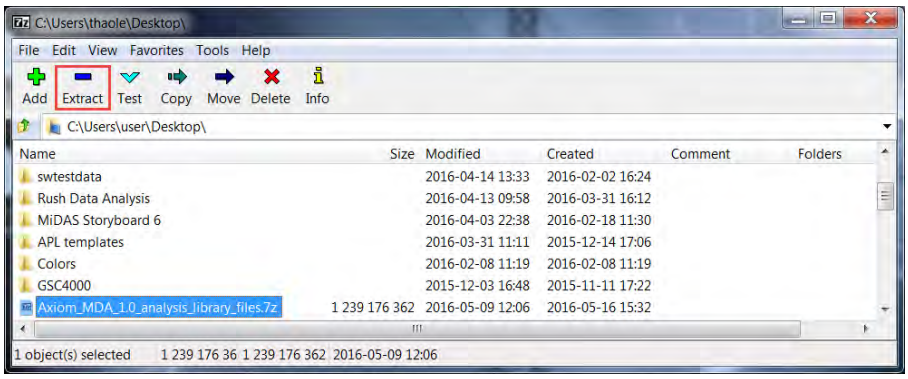


Figure 2.7 7-zip main window

The 7-zip Extract window appears (Figure 2.8).

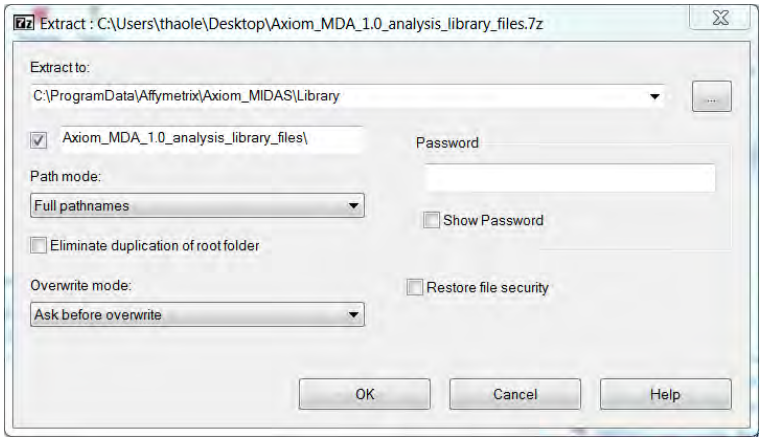


Figure 2.8 7-zip Extract window

7. In the **Extract to** field, enter C:\ProgramData\Affymetrix\Axiom_MIDAS\Library (as shown in Figure 2.8).
8. Click **OK**.

   The Axiom MiDAS library file is ready for use.

2 Performing an analysis

Setting up an analysis

1. From the main Axiom MiDAS **Setup** tab, click the **Add Files** button (Figure 2.9).

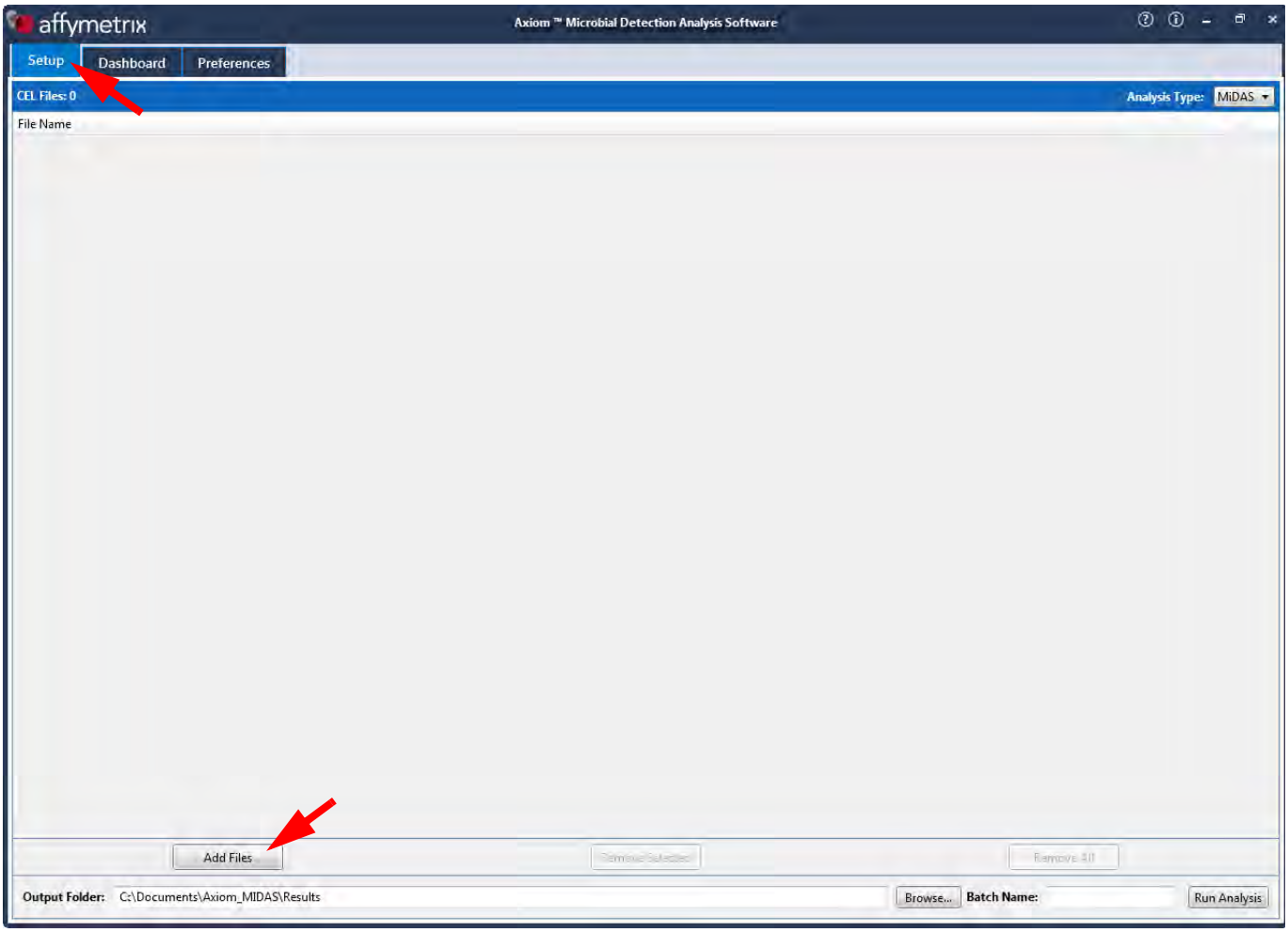


Figure 2.9 Main window

Importing CEL files

1. Click **Add Files**.

   The **Select CEL files** window appears.
2. Navigate to your CEL file location.
3. **Ctrl+click**, **Shift+click**, or **Ctrl+a** to select (highlight) the CEL files you want to analyze, then click **Open**.

   The **CEL Files** pane populates and displays your imported CEL files (Figure 2.10).

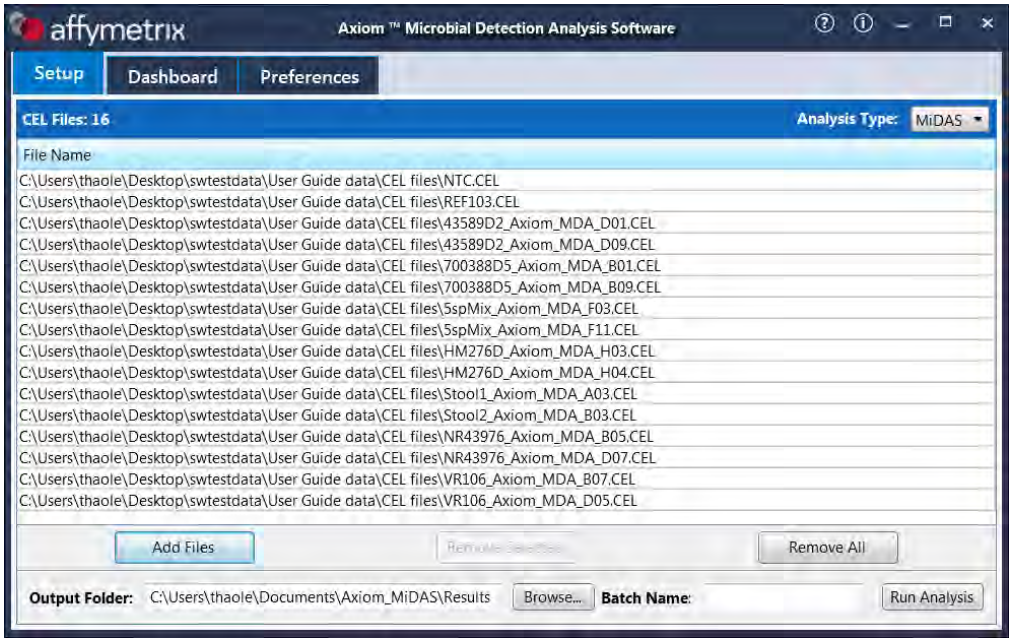


Figure 2.10 Populated CEL File pane example

To remove a CEL file(s):

1. Single-click, **Ctrl+click**, or **Shift+click** to highlight the files you want to remove.
2. Click Remove Selected.

To remove all CEL files:

1. Click Remove All.

Selecting an Analysis Type

1. Click the **Analysis Type** drop-down to select the type of analysis you want to run (Figure 2.10).

Figure 2.11 Analysis Type drop-down menu

- **MiDAS**: Select the Axiom MiDAS workflow to utilize the Axiom MiDAS algorithm to determine which micro-organisms likely comprise the samples that were run on Axiom™ Microbiome Arrays.
- **QC**: Select the QC workflow to generate the Dish QC (DQC) value for human control samples. This workflow produces a single QC Viewer window, as shown in Figure 2.17 on page 37.

Output Folder path

The output folder (Figure 2.12) is where your completed analysis results are saved (stored).

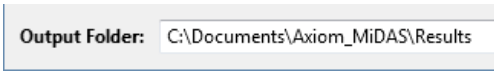


Figure 2.12 Example: Output Folder path field

Assigning a New Output Folder Path (Optional)

1. If you want to change the output folder path that is assigned by default, click Browse....

   An Explorer window appears.
2. Navigate to the output path folder location you want, then **Select Folder**.

   Your newly assigned output folder path is now displayed.

Adding sub-folders (optional)

Note: To better organize your output results, you can add sub-folders to your newly assigned output result path's folder.

To add sub-folders to your newly assigned result path's folder:

1. Click Browse... to return to your assigned output path and/or folder.
2. In the Explorer window, click **New Folder**.
3. Enter a sub-folder name.
4. Click **Select Folder**.

   The newly created sub-folder now appears in the output result information window.
5. Repeat the above steps 1-4 to add more sub-folders, then click **Select Folder**.

Assigning a Batch Name

1. Enter a name in the **Batch Name** field (Figure 2.12).

Figure 2.13 Enter a Batch Name

During the analysis process, a folder (with the same name as your entered batch name) is auto-generated. This folder includes all the necessary files needed to view your completed analysis results in the Axiom MiDAS Viewer.

Running your analysis

Note: To maximize the speed of your analysis run, make sure you close all other open and/or running programs before running your analysis.

1. Click Run Analysis.

   The Dashboard window tab appears and shows the status of your running analysis (Figure 2.12).

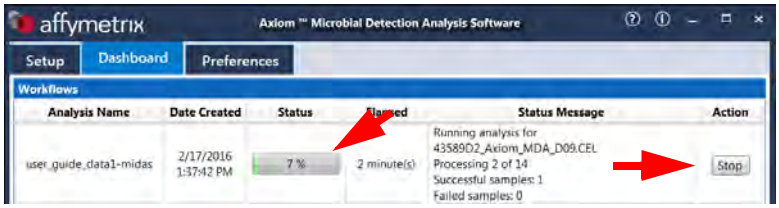


Figure 2.14 Dashboard window/tab - Status bar and Stop button example

2. Click Stop to cancel an analysis in progress.

Using the dashboard

The **Dashboard** tab displays your completed analysis results (Figure 2.15).

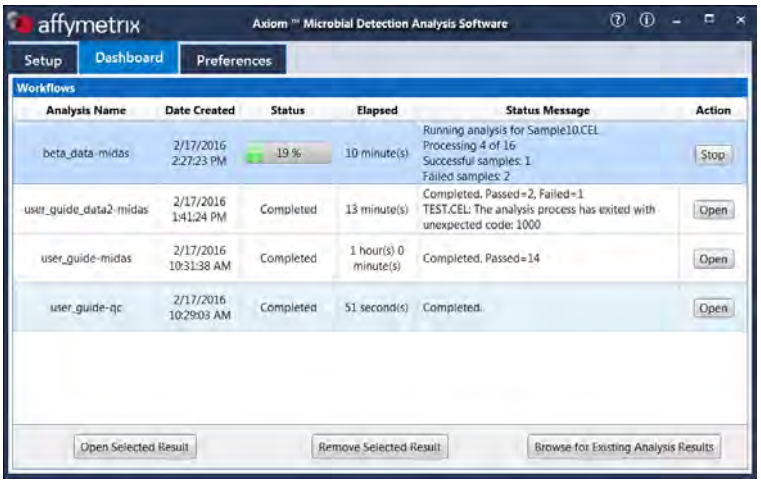


Figure 2.15 Dashboard window

Open selected result(s)

Do one of the following to open a selected result:

- Click Open (Figure 2.16).
- Single click a study, then click Open Selected Result.
- Double-click the completed analysis.
- Right-click the completed analysis, then click **Open**.

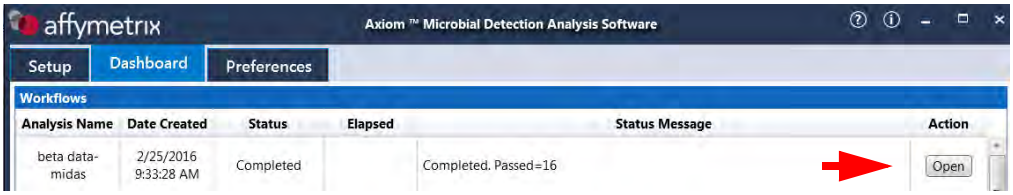


Figure 2.16 Completed analysis

The Axiom MiDAS Viewer appears. Please refer to Chapter 3, "Axiom™ MiDAS Viewer: Summary and Sample" for instructions on using the Axiom MiDAS Viewer.

Remove selected result(s)

Do one of the following to remove a selected result:

- Click to highlight the analysis you want to remove, then click Remove Selected Result.
- Right-click the highlighted analysis, then click **Remove from List**.

Browsing For existing analysis results

Do the following if a study is not listed on the Dashboard:

1. Click Browse for Existing Analysis Results.

   A **Select the results folder to open** window appears.
2. Navigate to, then click an existing analysis. Click **Select Folder**.

   Your selected analysis appears in the **Dashboard** tab.
3. Click Open.

   The Axiom MiDAS Viewer appears.

   Please refer to Chapter 3, "Axiom™ MiDAS Viewer: Summary and Sample" for instructions on using the Axiom MiDAS Viewer's **Summary** and **Sample** windows.

QC Workflow Analysis Viewer

1. After a successful QC Analysis run, click Open .
   The QC Viewer window appears (Figure 2.16).

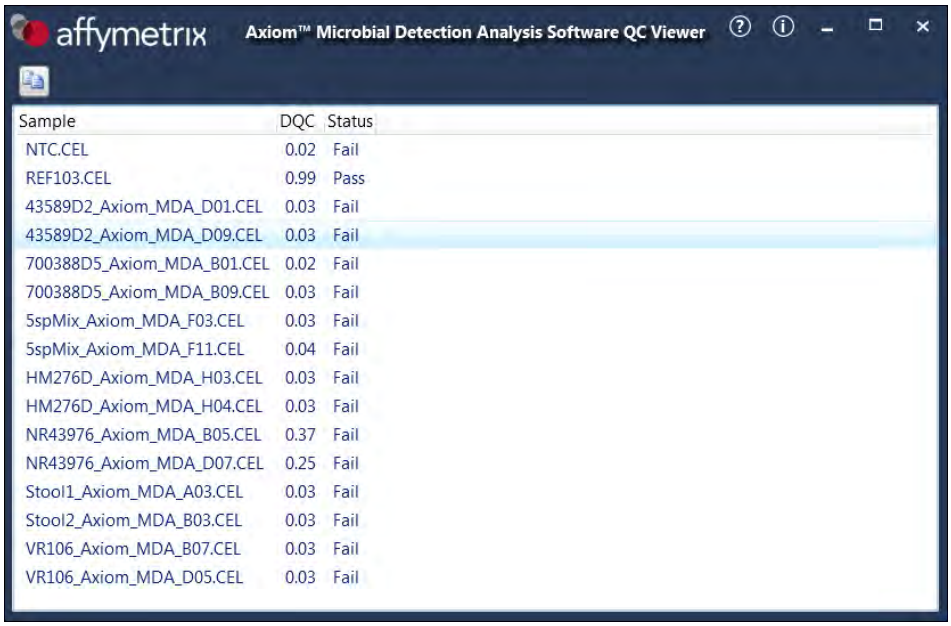


Figure 2.17 QC Viewer window

3 Axiom™ MiDAS Viewer: Summary and Sample

Summary

After setting up, running, then opening a successfully completed analysis, the Axiom™ MiDAS Viewer opens at the **Summary** tab (Figure 2.18). It displays a cross-sample snapshot of your analysis, including a list of samples analyzed, taxonomic richness values, and high level sample to sample comparisons.

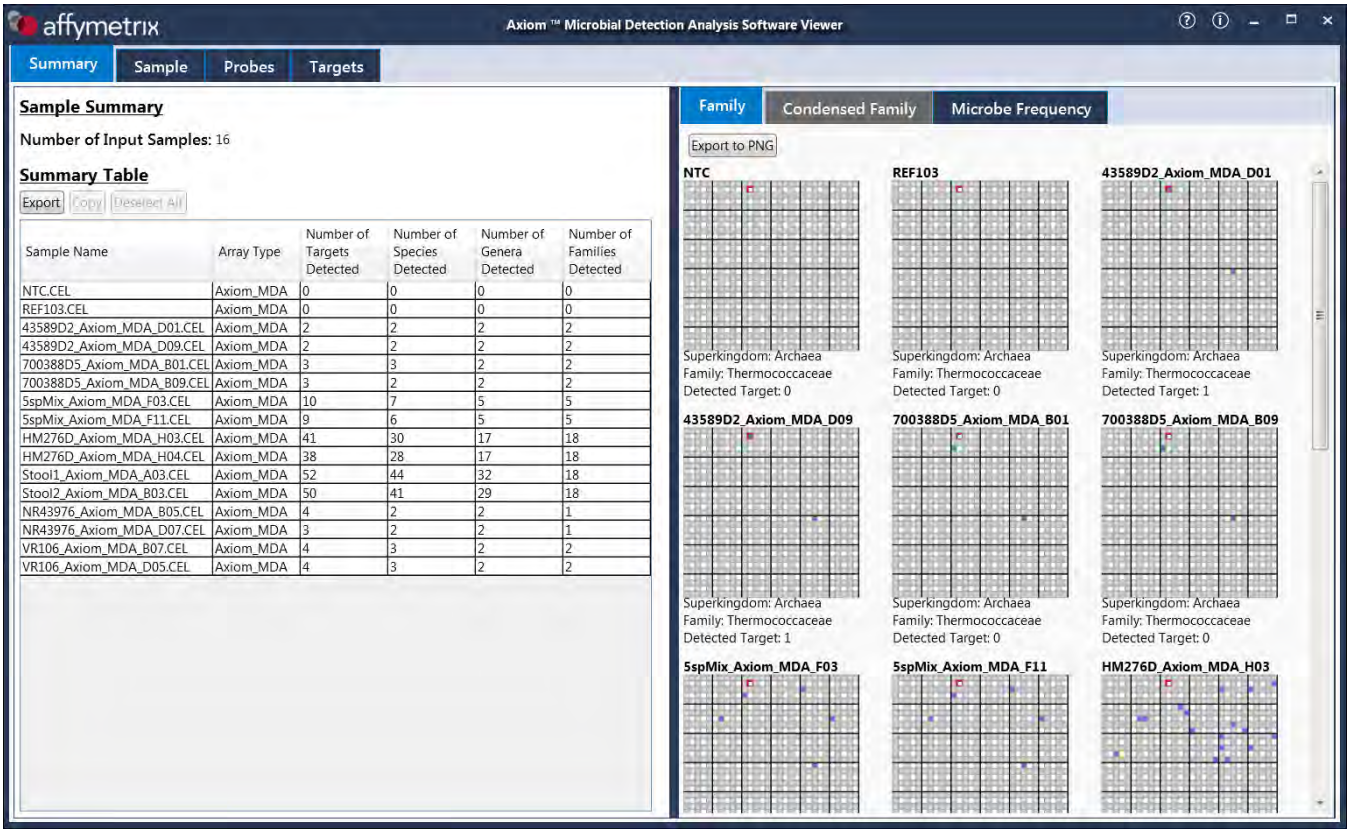


Figure 2.18 Summary tab

Using the Summary Table

Sample Summary

The Sample Summary displays your total sample count (Figure 2.19).

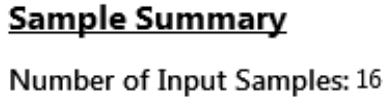


Figure 2.19 Sample Summary

Summary Table Column Definitions

Column name	Description
Sample Name	Name of the sample CEL file.
Array Type	Microarray type.
Number of Targets Detected	Number of targets in the algorithm model.
Number of Species Detected	Number of species in the algorithm model.
Number of Genera Detected	Number of genera in the algorithm model.
Number of Families Detected	Number of families in the algorithm model.

Note: The number of families shown here only represent targets with known family assignments from the Axiom™ MiDAS database.

Sorting Columns

1. Right-click a column header.

   A column function menu appears.

2. Click either **Sort by Ascending** or **Sort by Descending**.
   - The column is now sorted appropriately.
   - Click **Unsort** to return to a column's default sorting order.

Rearranging columns

1. Click a column you want to move.
2. Drag it (left or right) to its new location, then release the mouse button.

   The column is now in its new position and retained (auto-saved) for future sessions.

Copying Column Data to a Clipboard

1. **Ctrl+click**, **Shift+click**, or **Ctrl+a** to select (highlight) the samples you want to copy to a clipboard, then click the Copy button.

   The column data is now ready for pasting (**Ctrl+v**) onto a clipboard.

Exporting Summary Table Results

1. Click the Export button.

   The **Export to TXT** window appears.
2. Navigate to where you want to save the .TXT file, enter a filename, then click **Save**.

Using the Summary Viewers

The Summary window tab features three unique graphic representations of your analyzed sample data.

- Family Tiles
- Condensed Family Tiles
- Microbe Frequency

Family Tiles

Block of tiles depicting families detected by Axiom™ Microbiome Array. Each highlighted tile represents a single microbial family detected in the sample. When a tile is selected, a red outline and informational text appears. Each family is associated with a fixed tile position across all samples. In addition, there is a single added tile at the end of each superkingdom group to represent targets with an unassigned Family.

Note: The Family tiles (block views) displayed do not represent positions on the array.

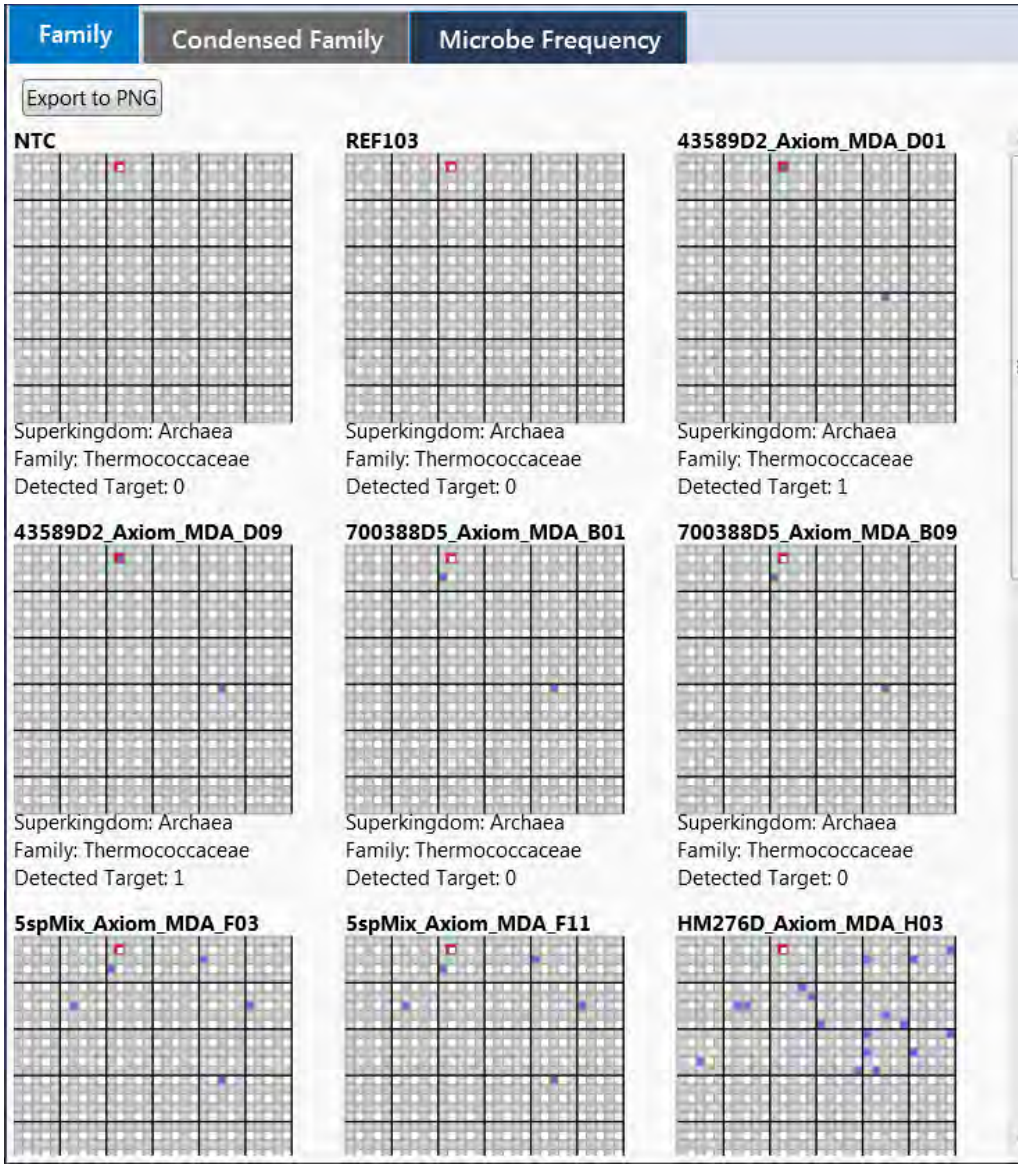


Figure 2.20 Family Tiles view

Using the Family Tile View

By default, all sample names listed in the Summary table are displayed (in tile form) within the **Family** tab. A maximum of 9 samples are displayed (Figure 2.20).

Note: Use the scroll bar and/or click and drag the center divider left to reveal all displayed tiles.

1. From the Summary table, **Ctrl+click**, **Shift+click**, or **Ctrl+a** to select (highlight) the samples you want to view in the **Family** tab.

   Your selected samples are now displayed in the **Family** tab, as shown in Figure 2.21.

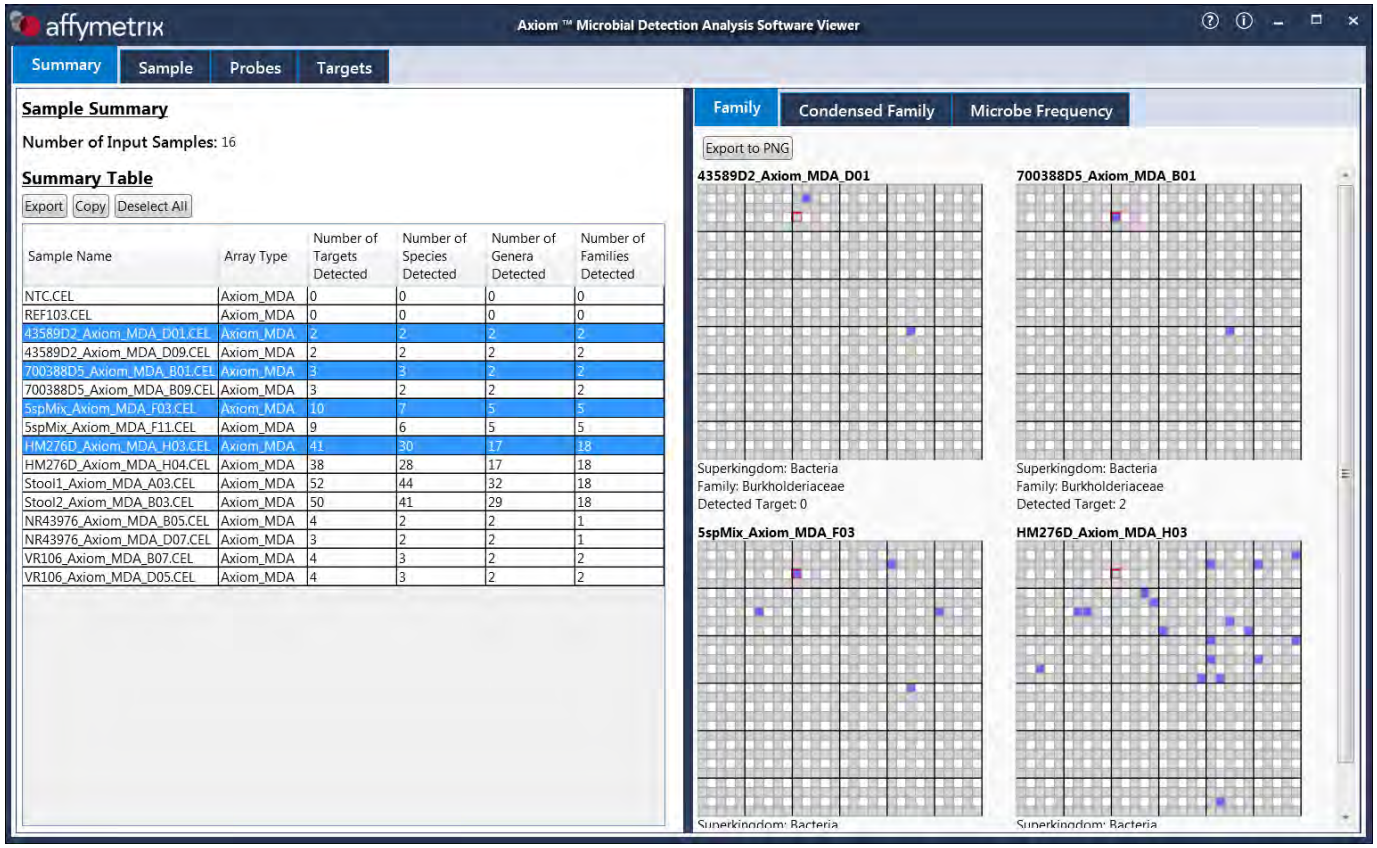


Figure 2.21 Example: Highlighted Summary Table entries to Family tiles

2. Click Deselect All to remove all highlighted (selected) samples from the Summary table.

To view a specific tile well-point

Note: The Family tile (block view) displayed does not represent positions on the array.

1. Click any tile of interest to reveal its details, as shown in Figure 2.22.

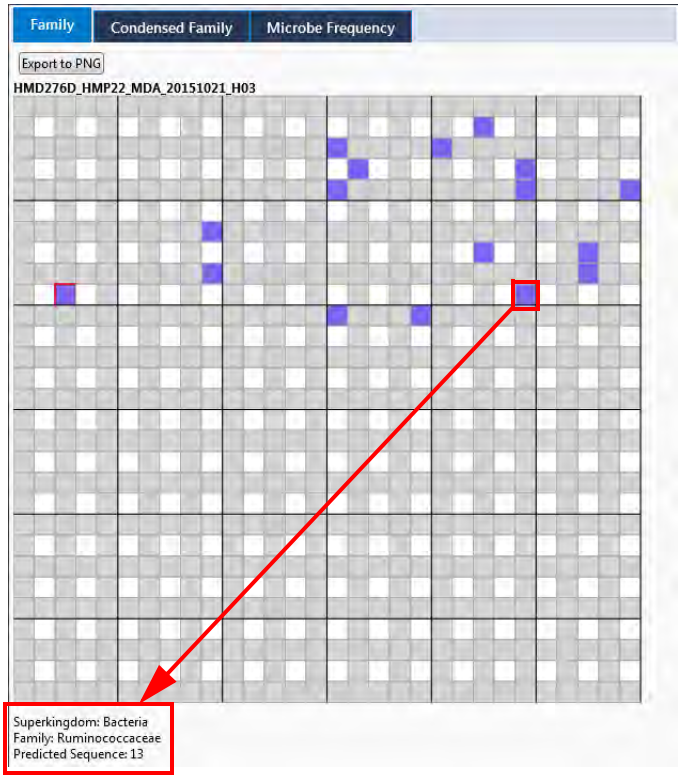


Figure 2.22 Well-point properties

Saving a Family Tile View to a PNG Graphic File

Note: Before saving your view as a graphic file, maximize it by dragging the center divider left.

1. Click the Export to PNG button.

   The **Export to PNG** window appears.

2. Navigate to where you want to save the .PNG file, enter a filename, then click **Save**.

Condensed Family Tiles

The **Condensed Family** tiles tab filters/condenses each Family tile into a compact view showing only Families detected in at least one sample.

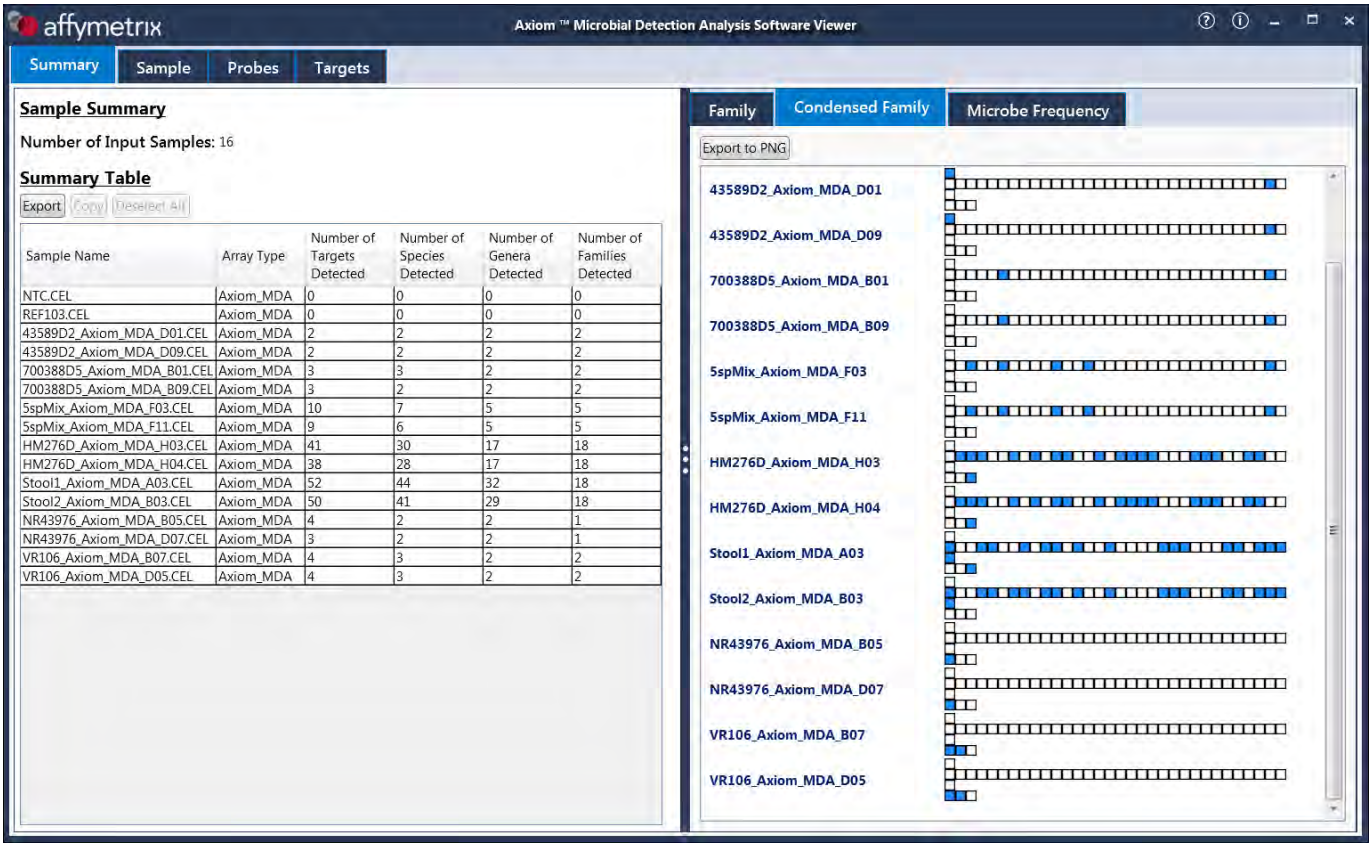


Figure 2.23 Condensed Family view

Using the Condensed Family Tiles View

1. From the Summary table, **Ctrl+click**, **Shift+click**, or **Ctrl+a** to select (highlight) the samples you want to view in the **Condensed Family** tab.

   The corresponding Condensed Family rows are now displayed, as shown in Figure 2.24.

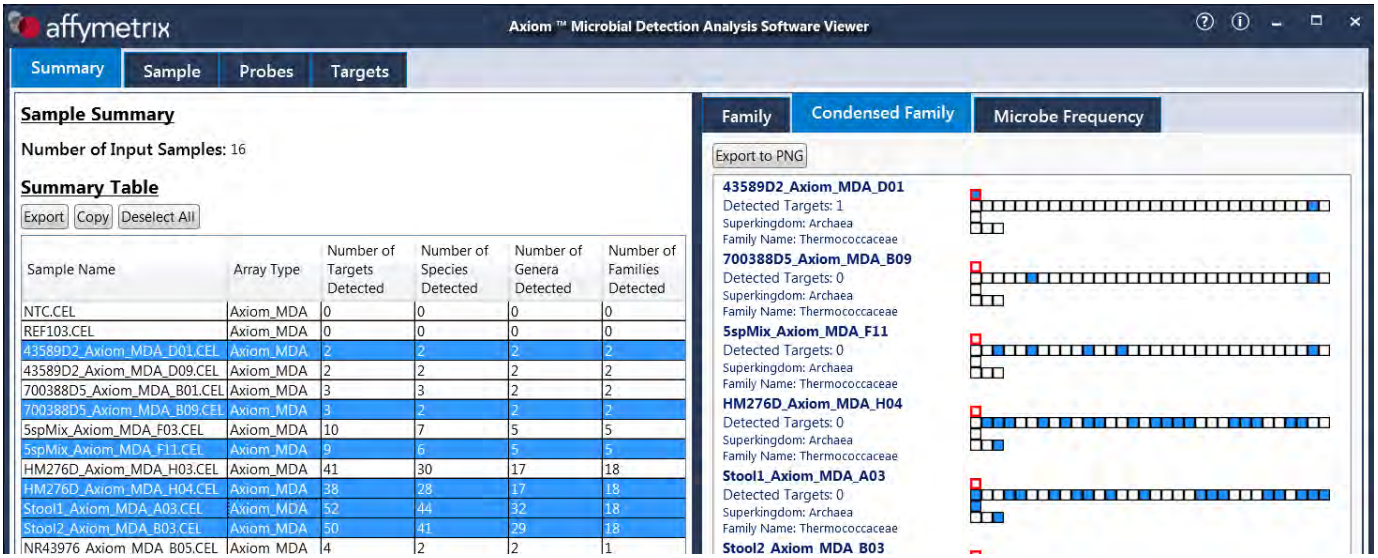


Figure 2.24 Example: Highlighted Summary Table entries to Condensed Family rows

To view a specific tile:

1. Click any family tile of interest, as shown in Figure 2.25.

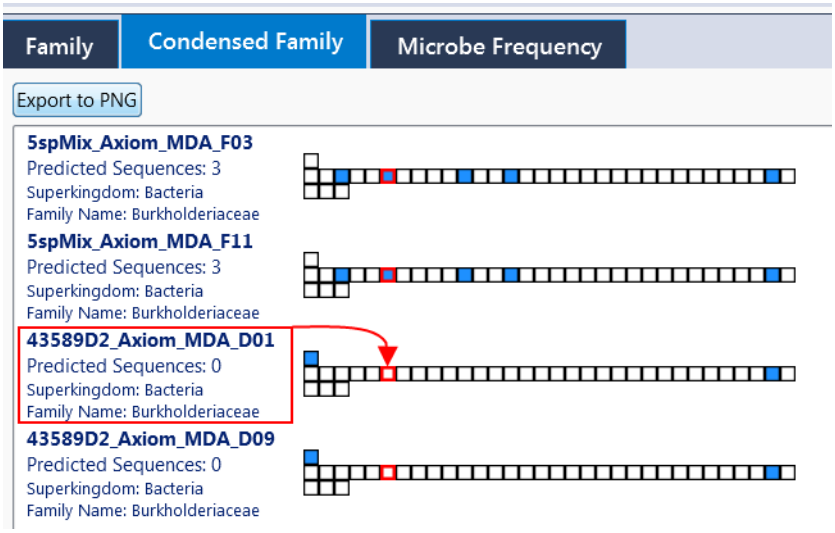


Figure 2.25 Detected Family properties

Saving a Condensed Family View to a PNG Graphic File

Note: Before saving your view as a graphic file, maximize it by dragging the center divider left.

1. Click the Export to PNG button.

   The **Export to PNG** window appears.

2. Navigate to where you want to save the .PNG file, enter a filename, then click **Save**.

Microbe Frequency

Microbe Frequency graph depicts the frequency with which known families or species of microbes appear in your sample set.

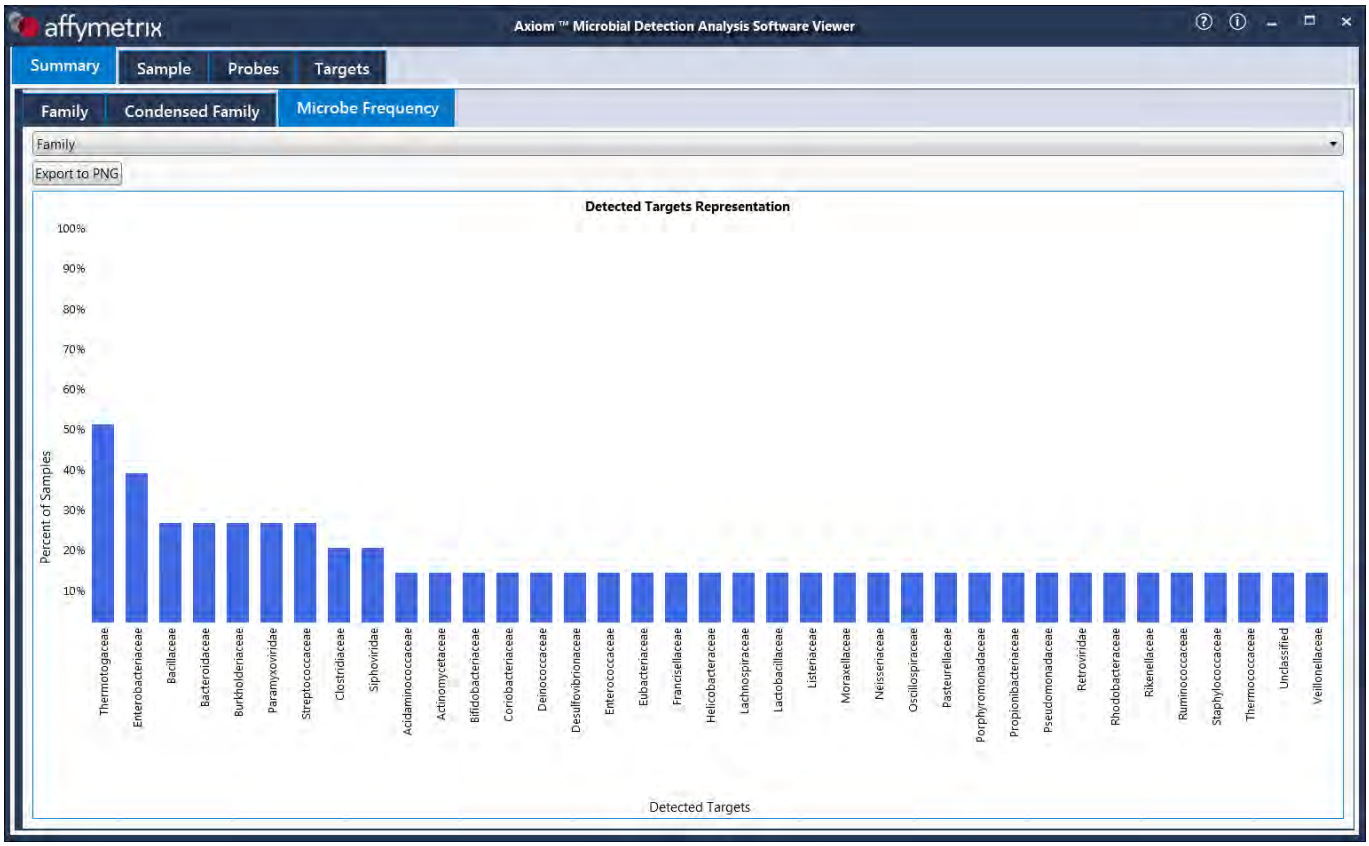


Figure 2.26 Microbe Frequency graph example

Using the Microbe Frequency Graph

1. At the graph, click or **Ctrl+click** a graph bar(s) of interest.

   The selected graph bar highlights the associated sample names within the Summary table, as shown in Figure 2.27.

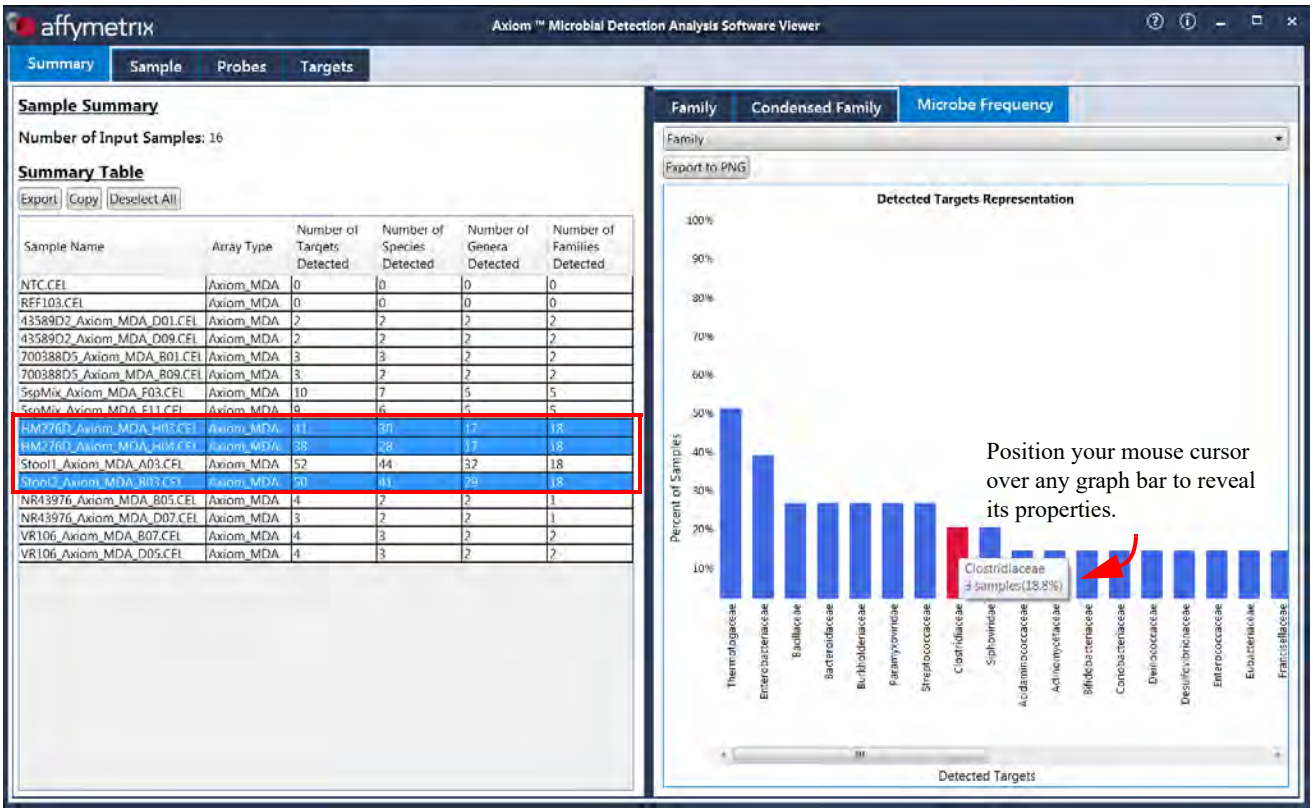


Figure 2.27 Example: Highlighted Microbe Frequency graph bar to Sample table data

2. To deselect highlighted table rows, click Deselect All . Alternatively, click outside the bar graphics to deselect a currently highlighted bar.

Changing the Microbe Graph View

1. By default, the graph is set to **Family**

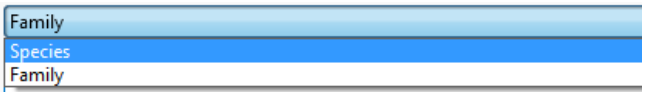


Figure 2.28 Microbe graph drop-down menu

2. Click the drop-down (Figure 2.28), then click **Species**.

   The graph is now set to **Species**.

Saving a Graph View to a PNG Graphic File

Note: Before saving your view as a graphic file, maximize it by dragging the center divider left.

1. Click the Export to PNG button.

   The **Export to PNG** window appears.

   Navigate to where you want to save the .PNG file, enter a filename, then click **Save**.

Sample

The **Sample** window tab (Figure 2.29) displays the list of organisms most likely to be in the sample. Its user-interface uses different parameters from the algorithm and is dynamic by allowing you to apply different thresholds to further investigate your analyzed data.

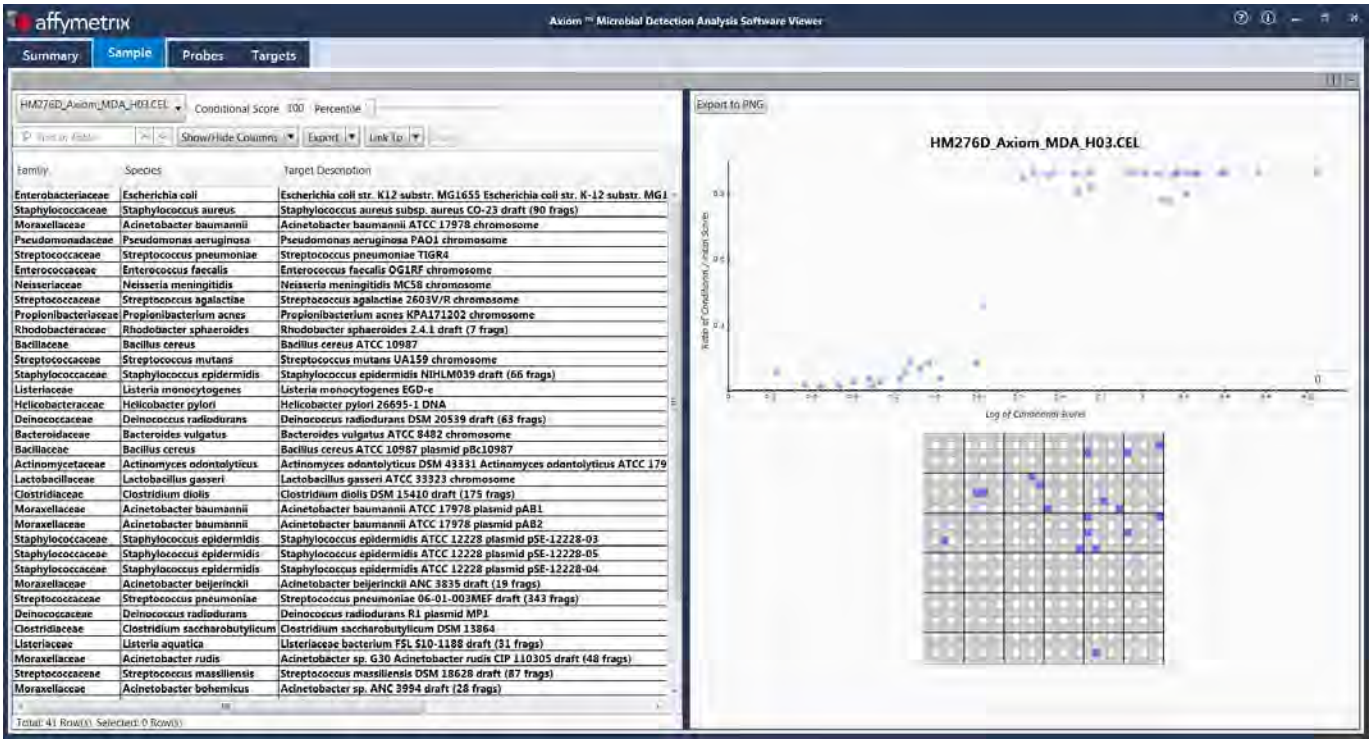


Figure 2.29 Sample window tab

Using the Sample Table

The currently displayed sample file and its properties reside in the Sample table.

The table entries shown in **bold** represent detected targets.

Split-Screen Options

By default, a side by side split-screen configuration is displayed.

To change to a top and bottom configuration:

1. Click the **Horizontal Split** icon (Figure 2.30).

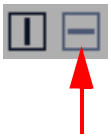


Figure 2.30 Horizontal Split icon

To view a different analyzed file:

1. Click the drop-down menu (Figure 2.31), then click to select the file you want to view in detail.

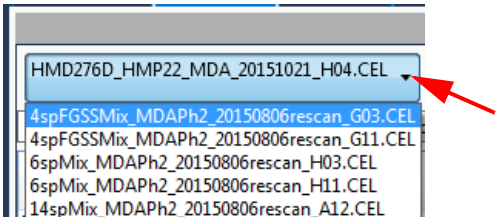


Figure 2.31 Sample File to view drop-down menu

After a few moments, your newly selected sample appears in the Sample table and adjacent graphic views.

Finding keywords

1. Click inside the Find in Table field, then type your keyword.
2. Click the buttons to start your keyword search.

Note: By default, the search tool locates matches that contain your (case insensitive) search inquiry. Use wild-card (*) characters to aid in your search. Example: ABC* = Any string that begins with "ABC". Use (") to search for exact content matches. Example: "ABC".

If a match is found, the first appropriate table row (below your selection) is highlighted. Use the buttons to view additional matches.

Showing or Hiding Columns

1. Click the **Show/Hide Columns** drop-down menu. (Figure 2.31)

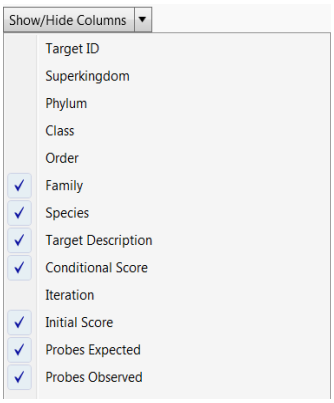


Table 2.3 Show/Hide Columns drop-down menu

2. Click a column's label to show it on the table. Click a check mark ✓ to hide the column from the table. See "Sample Table Column Definitions" on page 49 for their definitions.
3. Click outside the **Show/Hide Columns** drop-down menu to close it.

Sample Table Column Definitions

Column Name	Description
Target ID	Unique identifier for the target.
Superkingdom	Superkingdom classification of the detected target.
Phylum	Phylum classification of the detected target.
Class	Class classification of the detected target.
Order	Order classification of the detected target.
Family	Family classification of the detected target.
Species	Species classification of the detected target.
Target Description	Database record of the detected target.
Conditional Score	The conditional likelihood ratio for a target if detected targets from previous iterations are present.
Iteration	The analysis round in which the target was chosen.
Initial Score	The likelihood ratio for a target if no other targets are present.
Probes Expected	Number of distinct probes expected to give a positive signal for the detected target.
Probes Observed	Number of distinct probes giving a positive signal for the detected target.
Total Expected Probe Hyb Probability	Sum of the hybridization (detection) probabilities of distinct expected probes.
Total Observed Probe Hyb Probability	Sum of the hybridization (detection) probabilities of distinct observed probes.

Note: Detected probe counts shown on this table refer to those detected above threshold at unique genomic positions, please see Table 2.4 on page 57.

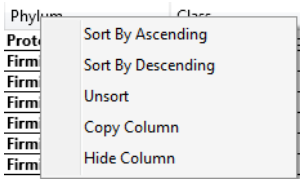
Rearranging Columns

1. Click a column you want to move.
2. Drag it (left or right) to its new location, then release the mouse button.

   The column is now in its new position and retained (auto-saved) for future sessions.

Sorting Columns

1. Right-click a column header.
   A column function menu appears.

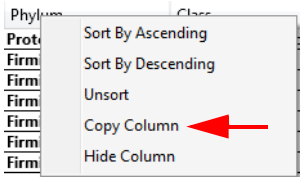


2. Click either **Sort by Ascending** or **Sort by Descending**.

   The column is now sorted appropriately.
3. Click **Unsort** to return to the default sorting order.

Copying Columns

1. Click to select a column you want to copy to a clipboard, then right-click it.
   A column function menu appears.

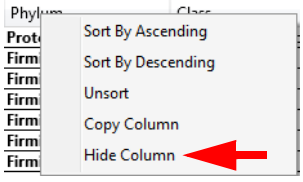


2. Click **Copy Column**.

   The column data is now ready for pasting (**Ctrl+v**) onto a clipboard.

Hiding Columns

1. Click to select a column you want to hide from the table, then right-click it.
   A column function menu appears.



2. Click **Hide Column**.

   The column is now hidden from the table.

Using the Sample Viewer

1. At the Sample table, **Ctrl+click**, **Shift+click**, or **Ctrl+a** to select (highlight) the targets you want to view in the plot graph. Alternatively, click the graph plots of interest to highlight their associated table entries.

   Your selections are highlighted in the plot graph and Family tile, as shown in Figure 2.32.

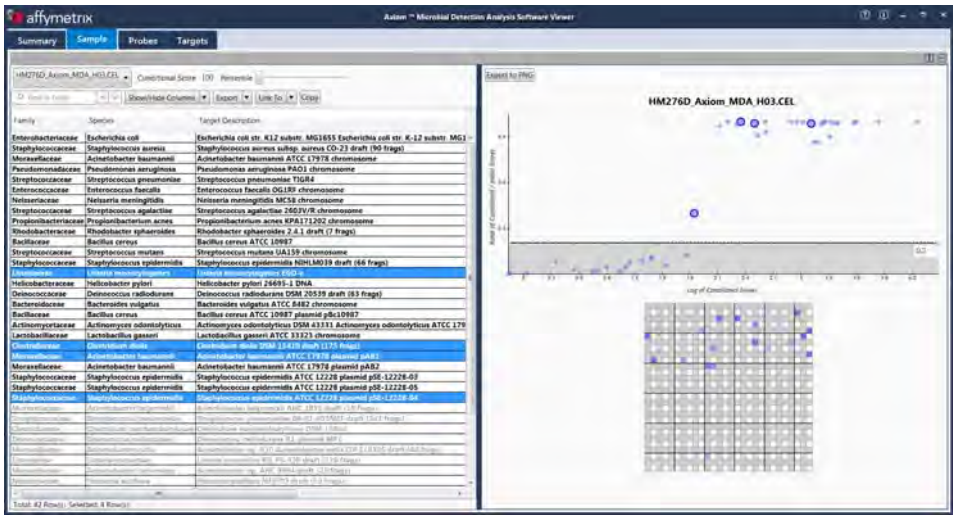


Figure 2.32 Highlighted Sample table entries to graph plots (horizontal view)

Conditional Score Slider

Enter a value into the text box or move the slider bar right to view secondary targets with a conditional score above your chosen percentile (Figure 2.33). Please refer to Chapter 4: "Initial and conditional scores" for more information on conditional scores.

Figure 2.33 Slider bar

Note: Secondary targets displayed in the Sample table appear as regular text, while detected targets are bold.

Moving the slider right shows the detected (blue) plots and any additional secondary (red) plots that fit your selection, as shown in Figure 2.34.

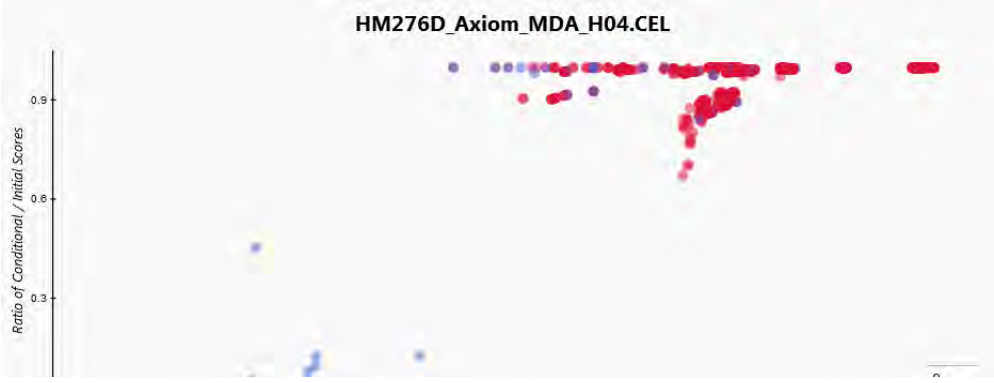


Figure 2.34 Red and blue graph plots

Setting a Threshold Value

1. Click the line on the graph (default set at 0.2).

   A double-headed arrow graphic appears.

2. Hold down the mouse button while dragging arrow graphic upwards to set your own threshold value, as shown in Figure 2.35.

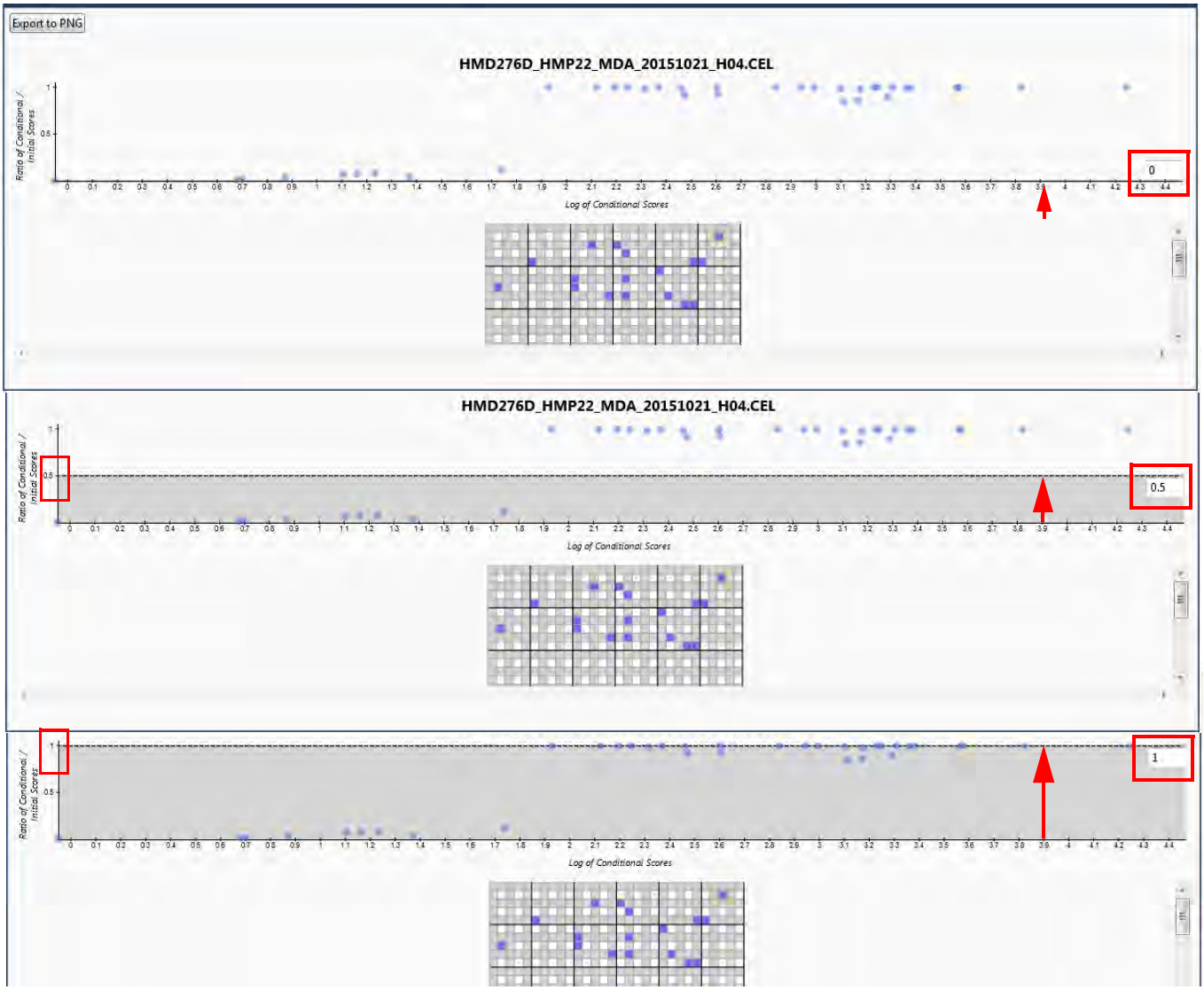


Figure 2.35 Highlighted Sample table entries to graph plots - Horizontal view example

Alternatively, click inside the threshold box 0, enter a new value, then press **Enter**.

Note: When a threshold is set above the 0.0 mark, all target data points *below* the line on the Scatter Plot are grayed out. These grayed out target data points can be viewed in the Sample Table.

Shading of the Family Tile auto-refreshes when a threshold is set above the 0.0 mark. In addition, all Family Tiles containing target data points *below* the threshold line are assigned a lighter highlighted color.

Accessing National Center for Biotechnology Information (NCBI)

This feature requires an active internet connection and the latest version of Internet Explorer.

1. Click to highlight a row entry for more information.
2. Click the Link To drop-down button.

   A menu appears listing the available external web sites that coincide with your selected row.

   All targets have links to at least one NCBI database:

   - **Bioproject**—Project driven database of biological data derived from a single organization or consortium.
   - **GenBank**—NIH database containing all publicly available annotated genetic sequences.
   - **Taxonomy**—Database of organism names and phylogenetic relationships.
3. Click to select the link you want to access.

   An internet browser window tab appears, as shown in Figure 2.36.

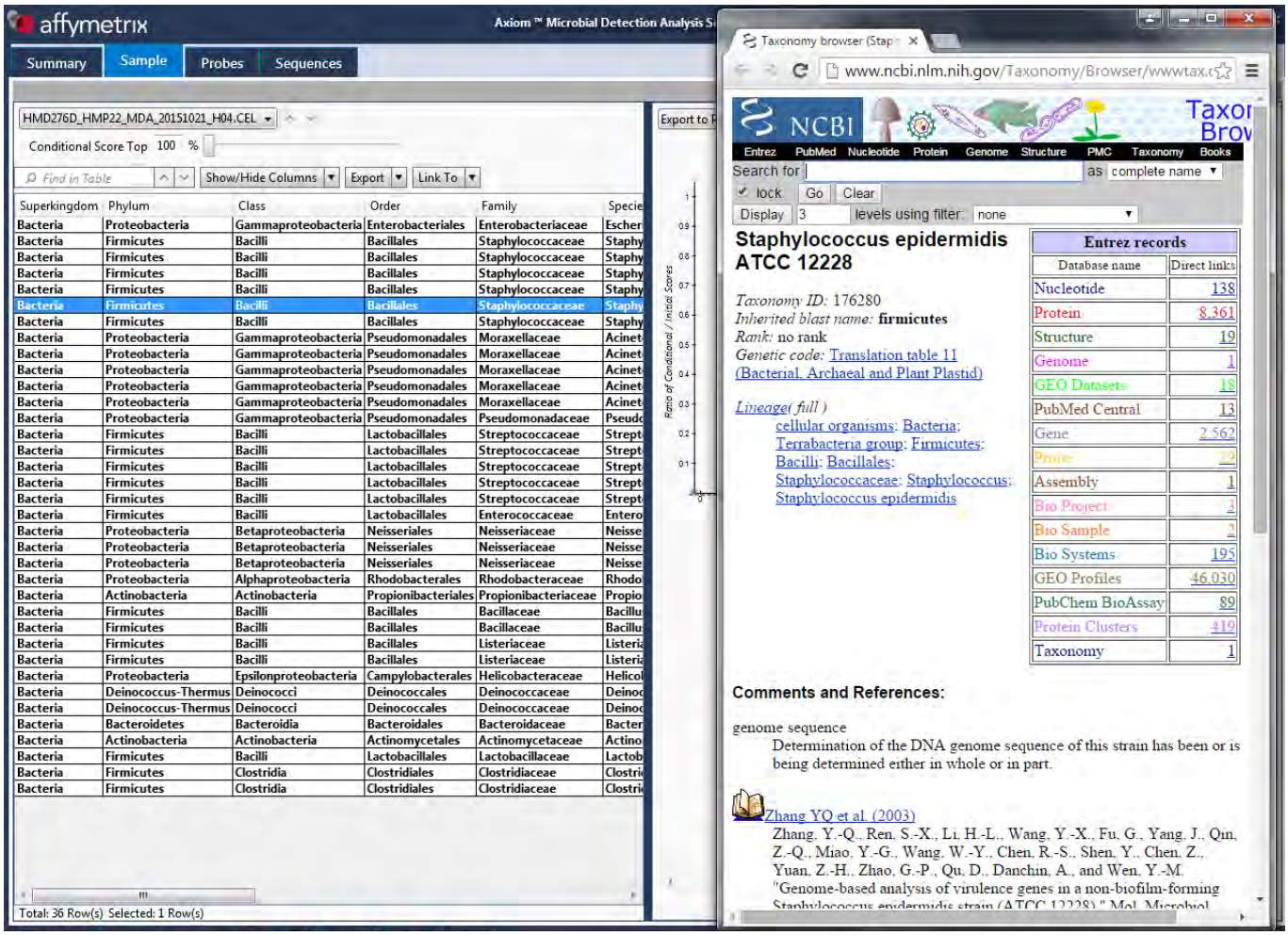


Figure 2.36 Sample window and NCBI tab example

Saving a Graph View to a PNG Graphic File

Note: Before saving your view as a graphic file, maximize it by dragging the center divider left.

1. Click the Export to PNG button.

   The **Export to PNG** window appears.

2. Navigate to where you want to save the .PNG file, enter a filename, then click **Save**.

Exporting Sample Table Results

1. Click the Export drop-down button.

   The following menu appears (Figure 2.37).

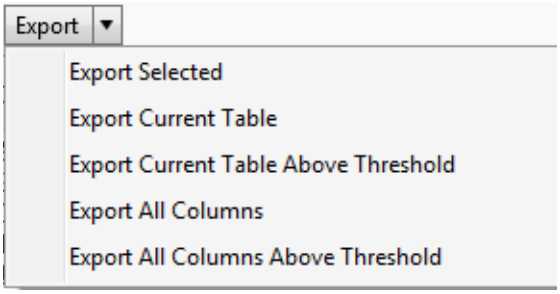


Figure 2.37 Exporting menu

2. Click to select one of the following:
   - **Export Selected**—Exports only your highlighted table rows and their displayed (shown) columns.
   - **Export Current Table**—Exports the table and its displayed (shown) columns.
   - **Export Current Table Above Threshold**—Exports entities above the set threshold with its displayed (shown) columns.
   - **Export All Columns**—Exports all hidden and shown table columns.
   - **Export All Columns Above Threshold**—Exports entities above the set threshold with all hidden and shown table columns.

   An **Export to TXT** window appears.

3. Navigate to a save location as you normally would, enter a filename, then click **Save**.

Please refer to Chapter 4, "Axiom™ MiDAS Viewer: Probes and Targets" for instructions on using the Axiom MiDAS Viewer's **Probes** and **Targets** windows.

4 Axiom™ MiDAS Viewer: Probes and Targets

Probes

The **Probes** tab displays probes for detected targets from the sample selected in the Sample view. Its table and adjacent plot graph provide data to support the targets detected by the Axiom™ Microbial Detection Analysis Software (MiDAS) application.

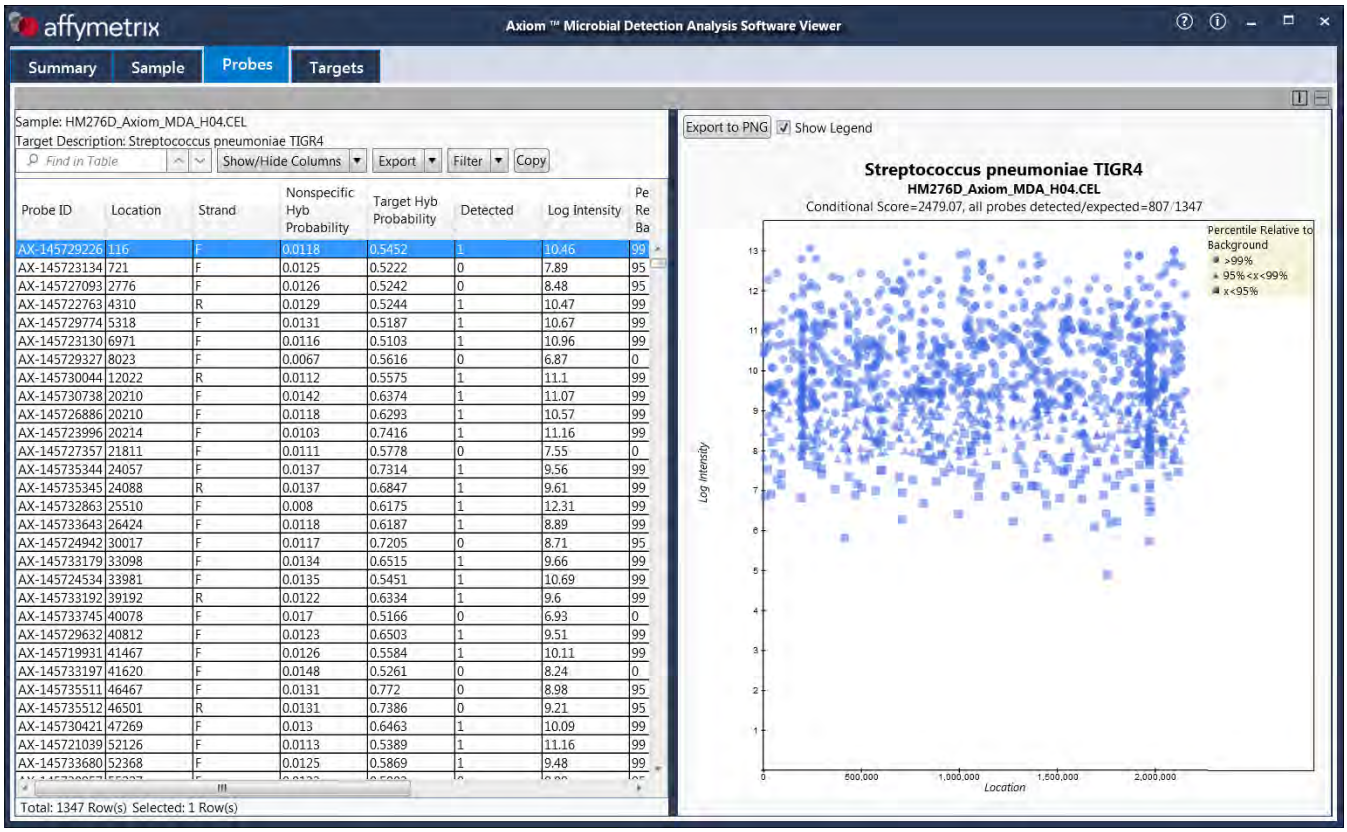


Figure 2.38 Probes Window tab

Split-screen options

By default, a side by side split-screen configuration is displayed.

To change to a top and bottom configuration:

1. Click the **Horizontal Split** icon (Figure 2.39).

Figure 2.39 Horizontal Split icon

Using the probes table

Finding keywords

1. Click inside the Find in Table field, then type your keyword.
2. Click the ^ v buttons to start your keyword search.

   Note: By default, the search tool locates matches that contain your (case insensitive) search inquiry. Use wild-card (*) characters to aid in your search. Example: ABC* = Any string that begins with "ABC". Use (") to search for exact content matches. Example: "ABC".

   If a match is found, the first appropriate table row (below your selection) is highlighted. Use the ^ v buttons to view additional matches.

Column headers

To show or hide table columns:

1. Click the **Show/Hide Columns** (Figure 2.40) drop-down menu.

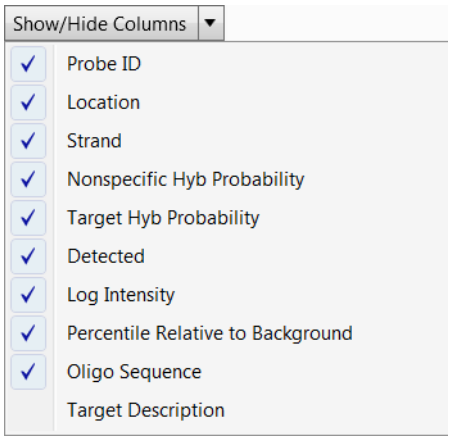


Figure 2.40 Show/Hide Columns drop-down menu

2. Click a column's label to show it on the table. Click a check mark ✓ to hide the column from the table. See "Probe table column definitions" on page 57 for their definitions.
3. Click outside the **Show/Hide Columns** drop-down menu to close it.

Probe table column definitions

Table 2.4 Column definitions

Column name	Description
Probe ID	Unique identifier for the probe.
Location	Genomic location of the selected target that the probe is expected to bind.
Strand	The binding orientation of the probe to the specified detected target.
Nonspecific Hyb Probability	Hybridization probability of the probe to a target other than the specified detected target.
Target Hyb Probability	Hybridization probability of the probe to the detected target.
Detected	Numerical representation of whether the probe was detected above threshold, with 0 being undetected and 1 being detected.
Log Intensity	Log of the raw intensity observed for the specified probe.
Percentile Relative to Background	Probe signal is observed to be above the specified percentile of random control probes.
Oligo Sequence	Nucleic acid sequence of the probe.
Target Description	Description of the detected target.

Rearranging columns

1. Click a column you want to move.
2. Drag it (left or right) to its new location, then release the mouse button.

   The column is now in its new position and retained (auto-saved) for future sessions.

Sorting columns

1. Right-click a column header.

   A column function menu appears. (Figure 2.40).

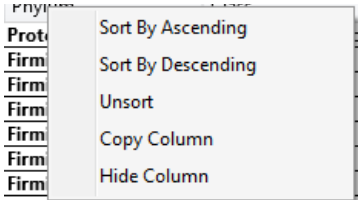


Table 2.5 Right-click column menu

2. Click either **Sort by Ascending** or **Sort by Descending**.
   The column is now sorted appropriately.
3. Click **Unsort** to return to the default sorting order.

Copying columns

1. Click to select a column you want to copy to a clipboard, then right-click it.
   A column function menu appears. (Figure 2.41)

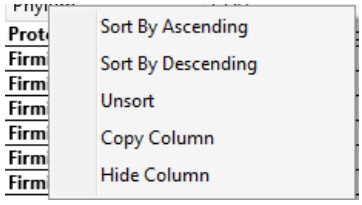


Figure 2.41 Right-click column menu

2. Click **Copy Column**.
   The column data is now ready for pasting (**Ctrl+v**) onto a clipboard.

Hiding columns

1. Click to select a column you want to hide from the table, then right-click it.
   A column function menu appears (Figure 2.42).

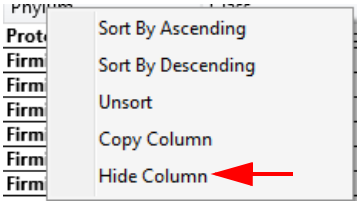


Figure 2.42 Right-click column menu

2. Click **Hide Column**.
   The column is now hidden from the table.

Filtering probe data

1. Click the Filter drop-down button.
   The following menu appears (Figure 2.43).

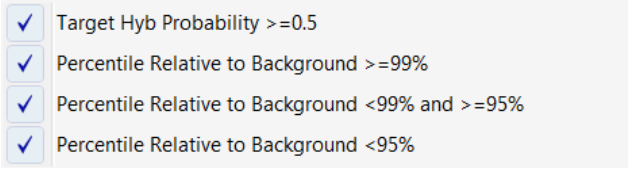


Figure 2.43 Filtering menu

Note: Probes with ≥0.5 target hybridization probabilities and signal ≥99% of background probes contribute positively to the initial score, while probes with ≥0.5 target hybridization probabilities and signal <99% of background probes contribute negatively to the initial score. Probes with <0.5 target hybridization probabilities are not considered in the initial score calculations.

2. Click a filter's label to show it on the table. Click a check mark ✓ to remove the assigned filter.

3. See the graph's legend (upper right) (Figure 2.44) to view the data plots affected by the assigned filters.
4. **Ctrl+click** probe points to reveal the specific filtering properties.

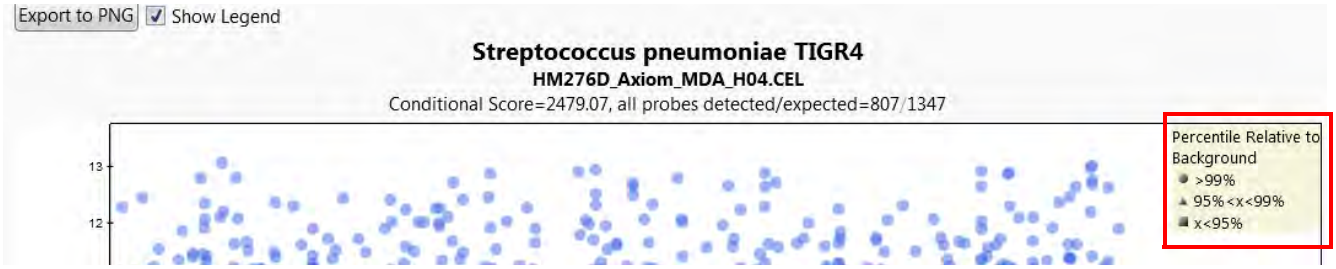


Figure 2.44 Filter legend

Using the Probe Viewer

1. Confirm your selection of **Sample** and **Target Description** (top left), as shown in Figure 2.45.

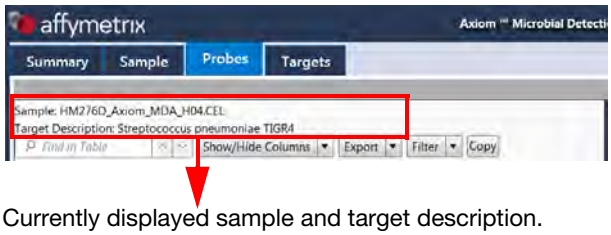


Figure 2.45 Sample and Target description in the **Probes** tab

2. At the **Probes** table, **Ctrl+click**, **Shift+click**, or **Ctrl+a** to select (highlight) the Probe IDs you want to view in the Probes graph. Alternatively, click the probe plots of interest to highlight their associated table entries. Your selections are highlighted, as shown in Figure 2.46.

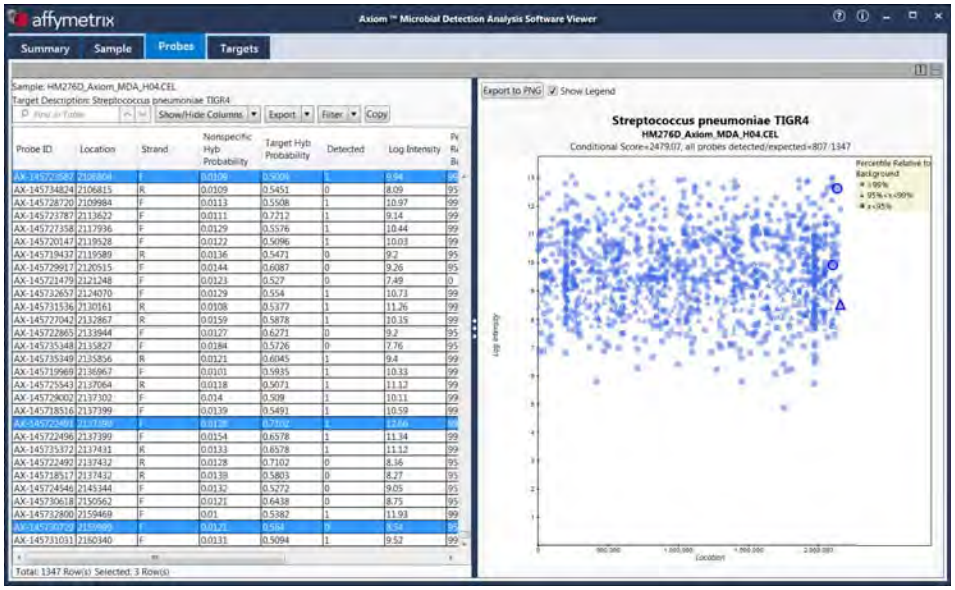


Figure 2.46 Example: Highlighted probe table entries displayed in plot graph

Viewing a different analyzed target

Use this method to view a different detected target within the same sample.

1. Click the Sample tab.
2. At the Sample table, click to highlight the detected target you want to view in the Probes window.
3. Click the Probes tab.

   Your newly selected detected target and its properties are now displayed in the Probes table and adjacent plot graph.

Exporting probe table results

1. Click the Export drop-down button.

   The following menu appears (Figure 2.47).

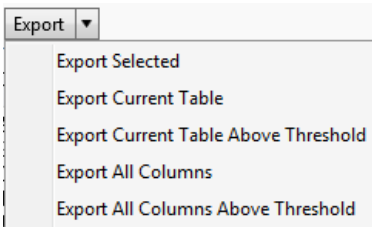


Figure 2.47 Exporting menu

2. Click to select:
   - **Export Selected**—Exports only highlighted table rows and their displayed (shown) columns.
   - **Export Current Table**—Exports the table and its displayed (shown) columns.
   - **Export Current Table Above Threshold**—Exports entities above the set threshold with its displayed (shown) columns.
   - **Export All Columns**—Exports all hidden and shown table columns.
   - **Export All Columns Above Threshold**—Exports entities above the set threshold with all hidden and shown table columns.
3. Navigate to a save location as you normally would, enter a **Filename**, then click **Save**.

Targets

The **Targets** tab displays all target calls across all your analyzed samples. Its table format allows you to quickly search for a target of interest and export your data for use in other applications.

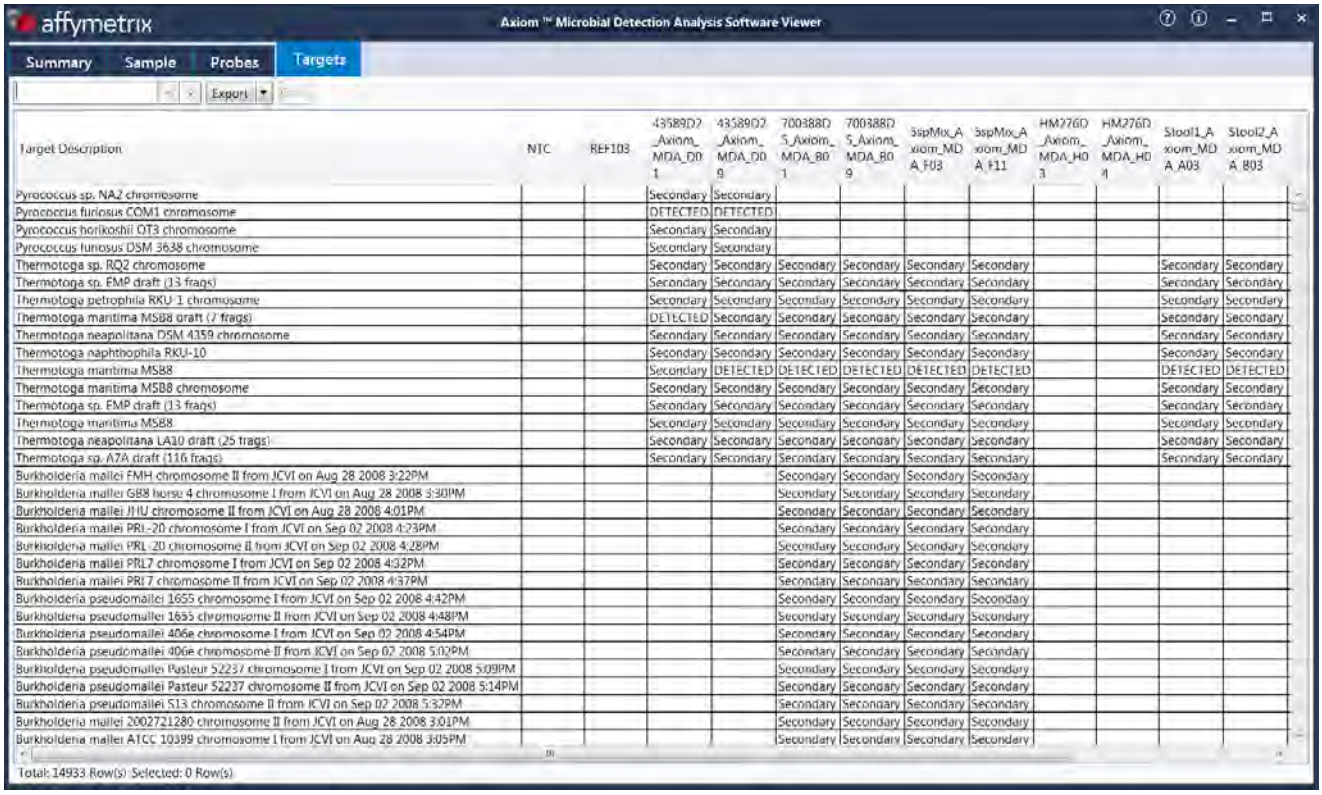


Figure 2.48 Targets window tab

Using the targets table

Finding keywords

1. Click inside the Find in Table field, then type your keyword.
2. Click the ▲ ▼ buttons to start your keyword search.

   Note: By default, the search tool locates matches that contain your (case insensitive) search inquiry. Use wild-card (*) characters to aid in your search. Example: ABC* = Any string that begins with "ABC". Use (") to search for exact content matches. Example: "ABC".

   If a match is found, the first appropriate table row (below your selection) is highlighted. Use the ▲ ▼ buttons to view additional matches.

Rearranging columns

1. Click a column you want to move.
2. Drag it (left or right) to its new location, then release the mouse button.

   The column is now in its new position and retained (auto-saved) for future sessions.

Sorting columns

1. Right-click a column header.
   A column function menu appears (Figure 2.49).

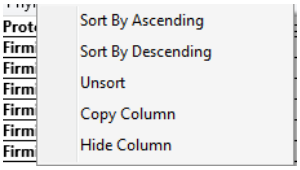


Figure 2.49 Right-click column menu

2. Click either **Sort by Ascending** or **Sort by Descending**.
   The column is now sorted appropriately.
3. Click **Unsort** to return to the default sorting order.

Method 1: Copying columns

1. Click to select a column you want to copy to a clipboard, then right-click it.
   A column function menu appears (Figure 2.50).

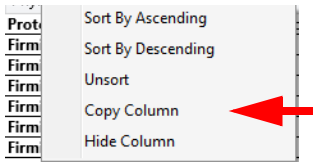


Figure 2.50 Right-click column menu

2. Click **Copy Column**.
   The column data is now ready for pasting (**Ctrl+v**) onto a clipboard.

Method 2: Copying columns

1. **Ctrl+click**, **Shift+click**, or **Ctrl+a** to select (highlight) the rows to copy.
2. Click the Copy button.
   The column data is now ready for pasting (**Ctrl+v**) onto a clipboard.

Hiding columns

1. Click to select a column you want to hide from the table, then right-click it.
   A column function menu appears (Figure 2.51).

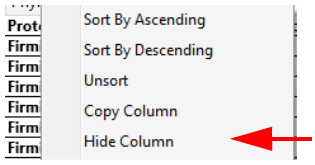


Figure 2.51 Right-click column menu

2. Click **Hide Column**.
   The column is now hidden from the table.

Exporting targets table results

1. Click the Export drop-down button.

   The following menu appears (Figure 2.52).

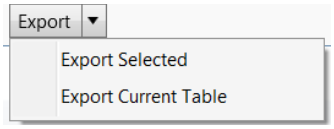


Figure 2.52 Exporting menu

2. Click to select:
   - **Export Selected**—Exports only your highlighted table rows and their displayed (shown) columns.
   - **Export Current Table**—Exports the table and its displayed (shown) columns.

Section 3:
Axiom™ Microbiome Array Data Analysis

1 Introduction

About this Section

This section provides information and instructions for using Axiom™ Microbial Detection Analysis Software (MiDAS) to analyze Axiom™ Microbiome Array data. It includes guidance on the interpretation of quality control (QC) analysis for plates and the results generated by the Axiom MiDAS algorithm.

Prerequisites

This section is intended for scientists, technicians, and bioinformaticians who need to analyze Axiom Microbiome Array data. This guide uses conventions and terminology that assume a working knowledge of bioinformatics, microarrays, quality control, and microbiome data analysis.

2 Microbial detection using Axiom™ MiDAS software

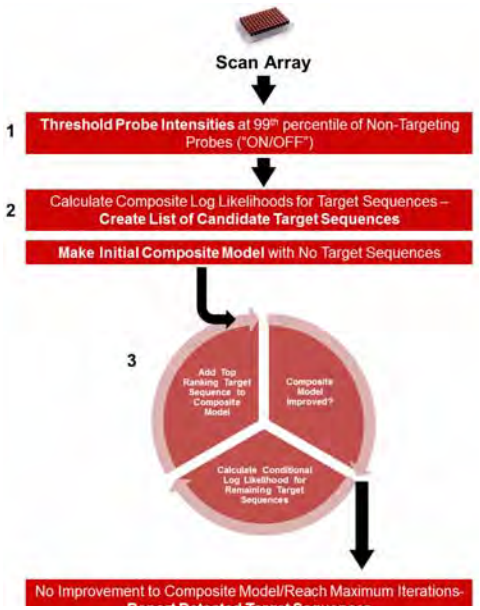
Introduction

The probes on Axiom™ Microbiome Array were designed to interrogate all organisms in Axiom™ MiDAS target sequence database. Some key definitions:

- **Probe**: an individual sequence on the chip, typically a 30-mer (e.g., atgctatcgtatacgtctgatagctagctagctag)
- **Target** or **target sequence**: a sequence record in the database typically a whole genome, but possibly a plasmid, or a composite draft genome (e.g., *Listeria monocytogenes* EGD-e or *Acinetobacter baumannii* ATCC 17978 plasmid pAB1). Most target sequences correspond to a single public domain database accession record. Probes were designed for both family level conservation (common to many sequences within a family) and individual target sequence identification (unique within each sequence). Detection probabilities are determined for all probes across all target sequences, so an individual probe may contribute to the identification of multiple targets. A target sequence represents an entry in the Axiom MiDAS database, which consists of microbial entries from public domain sequence repositories (e.g., NCBI).

The underlying algorithm used by Axiom MiDAS is based upon the CLiMax algorithm developed by Lawrence Livermore National Laboratory to analyze data from microbiome arrays. This algorithm detects target sequences within a sample by building a composite likelihood model that best explains the pattern of probe intensities observed on the array. Each term in the model is a target sequence whose likely presence in the sample helps explain a portion of the observed pattern of probe intensities. A likelihood score is calculated for each target sequence from probe-

specific factors.



Three stages of Axiom Microbial Detection Analysis Software.

- **Stage 1**: Thresholding.
- **Stage 2**: Initial target sequence identification.
- **Stage 3**: Iterations to improve model.

Axiom™ MiDAS algorithm stages

The Axiom MiDAS algorithm has 3 stages:

Stage 1: Thresholding

In stage 1, the threshold intensity for a probe is established using the intensities for the multiple negative control probes having similar length and %GC content. The probe is considered "on" or above threshold if its intensity is at or above the 99th percentile compared to those of all its corresponding negative control probes. At or above the 99th percentile threshold, there is only a 1% (or lower) probability that the observed intensity for the probe is due to nonspecific hybridization.

Stage 2: Initial target sequence identification

In stage 2, the algorithm identifies a set of candidate target sequences consisting of all those meeting **both** of the following criteria:

- a minimum of 8 corresponding probes that are "on" or above threshold
- a minimum of 20% of all their corresponding probes that are "on" or above threshold

Stage 3: Iteratively add target sequences to improve the model

In stage 3, the algorithm iteratively builds the final detected target sequence set. Initially, the detected set is empty, then a greedy procedure is used to find the target sequences that best explain the observed probe intensity pattern by maximizing the composite likelihood function. For more detailed information, please see Gardner *et al*.[1]

- In the first iteration, the algorithm computes an **initial** log composite likelihood ratio (log CL score) for each target sequence in the candidate target sequence set. The initial score for each target sequence is the ratio of the composite likelihood

[1]Gardner SN, Jaing CJ, McLoughlin KS, Slezak TR. A microbial detection array (MDA) for viral and bacterial detection. *BMC Genomics*. 2010 Nov 25;11:668. doi: 10.1186/1471-2164-11-668.

for a model in which only that target sequence is present to the likelihood of a model where **no** target sequences are present (i.e., one where all the "on" probe intensities are due to nonspecific hybridization). The composite likelihood for each target sequence specific model is computed by multiplying the detection probabilities for the positive probes and the non-detection probabilities for the negative probes (for the equation, please refer to "Calculating the initial composite likelihood" on page 73). The composite likelihood for the no-target sequences model is computed in a similar way, but using the nonspecific detection and non-detection probabilities instead.

- After computing initial log CL scores for each target, the algorithm chooses the target sequence with the highest score, then removes it from the candidate target sequence set and adds it to the detected target sequence set. Target sequences that had negative log CL scores in the first iteration are also eliminated from the candidate target sequence set. If the algorithm fails to find an initial target sequence with a positive log CL score, a model with no detected target sequences is returned.

In all subsequent iterations, the algorithm:

1. Calculates a **conditional** log CL score for each remaining candidate target sequence. The conditional log CL score for target sequence T in a given iteration is the ratio between the composite likelihoods for two models: a model where T is present together with the target sequences in the current detected target sequence set, and a model containing only the detected target sequence set (for the equation, please refer to "Calculating the conditional likelihood" on page 74).
2. Identifies the target sequence with the highest conditional score.
3. Removes it from the candidate target sequence set and adds it to the detected target sequence set.
4. Purges any candidate target sequences with negative conditional log CL scores.

The conditional score for a given target sequence indicates how well the addition of the target sequence to the model improves the model's explanation of the data. With each iteration, conditional scores of detected target sequences will decrease or remain the same relative to the initial score. The ratio of a target sequence's conditional score to its initial score indicates the amount of probe sequence similarity between the target sequence and other target sequences detected in previous iterations. In most cases, including two highly similar strains of an organism to the model provides little improvement over a model containing only one of the strains, thus a method to threshold on this ratio value and identify these similar strains is provided.

Iterations of the algorithm continue either until all remaining candidate target sequences have negative conditional scores, or until a maximum on the number of iterations is reached.

Secondary hits

Many of the targets in the database are very similar to other targets in the database. This means that the set of probes that describe one target are quite likely to describe a related target. This is handled in two ways:

1. The other targets (secondary hits) from the same family are listed below the called target from each iteration (primary hit) in the sample window in Axiom MiDAS Viewer. The secondary hits are listed in order of conditional likelihood (Figure 3.1).

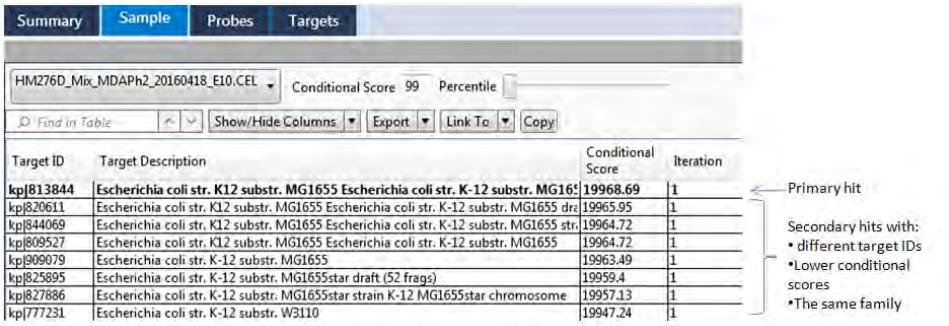


Figure 3.1 Sample table with primary and secondary hits

2. The probes that match the primary hit target are adjusted so that they can no longer provide as much support to any other target. This means that a secondary hit that shares almost all of the same probes as the primary hit will probably never come up as a positive hit of its own.

Target sequence probe detection probabilities

Target sequence specific detection probabilities (P(Y|X=1) in the equation (please refer to "Calculating the conditional likelihood" on page 74) are estimated using a generalized additive model that treats the log-odds of a probe binding to a specific target sequence as a smooth function of several experimentally identified predictor variables. These predictor variables include:

- BLAST similarity score between the probe and target sequences
- %GC content of the aligned probe and target sequences
- number and position of BLAST alignment mismatches between the probe and the target sequences
- number of GC bases in the upstream target sequence that is bound by random 9-mers in the second stage of the Axiom chemistry

The coefficients in this model were obtained by fitting it to data from a prototype array that were hybridized to samples comprised of known genome sequences.

Probabilities for **non-specific** binding (P(Y|X=0) in the equation (please refer to "Calculating the conditional likelihood" on page 74) are also estimated for each probe, using a different generalized additive model. In this model the predictor variables are:

- probe length
- %GC content of the probe
- entropy of the probe's trimer frequency distribution (a measure of sequence complexity)

Axiom™ MiDAS target sequence database

Axiom MiDAS target sequence database is comprised of select entries exported from Lawrence Livermore National Laboratory's KPATH pathogen surveillance pipeline and database (https://str.llnl.gov/str/April04/Slezak.html). The KPATH pipeline automatically and regularly checks public domain sequence repositories (e.g., NCBI) for new and relevant sequence entries and incorporates them into the database. See the following table for target categories represented in the Axiom MiDAS database. Axiom MiDAS database contains KPATH records that were current as of October 2014.

Domain	Number of families	Number of species	Number of target sequences
Archaea	31	370	606
Bacteria	278	6,901	34,254
Fungi	121	381	658
Protozoa	30	91	229
Virus	100	4,770	99,808
Total	**560**	**12,513**	**135,555**

Automated processing helps ensure that Axiom MiDAS and KPATH databases are comprehensive, but one of the expected trade-offs from inclusiveness is the incorporation of sequences that may add complexity to the interpretation of Axiom MiDAS results. Some types of target sequences that may be problematic when detected

by the Axiom MiDAS software include:

- **Redundant target sequences**: a given target sequence in Axiom MiDAS may be identical to other Axiom MiDAS target sequences, and both may appear in the detection results.
- **Draft genome sequences**: whole genome shotgun sequencing yields multiple contigs, and the Axiom MiDAS algorithm will occasionally give preference to these over finished sequences when making detection calls.
- **Sequences where the operational taxonomic units are enclosed in square brackets**: in some instances the detected target sequences may have operational taxonomic units (family, genus, species) enclosed in square brackets (e.g., *[Clostridium] stercorarium*). The square brackets represent recommended taxonomies (mainly based on genome trees), but not verified taxonomies (see: https://groups.google.com/forum/#!topic/qiime-forum/WnhQiB5q9Hc).

Redundant target sequences

A given target sequence in Axiom MiDAS may be identical to other Axiom MiDAS target sequences, and both may appear in the detection results with the same conditional score, initial score, probes expected, probes observed:

Family	Species	Target Description	Conditional Score	Initial Score	Probes Expected	Probes Observed
Enterobacteriaceae	Yersinia pestis	Yersinia pestis biovar Orientalis str. MG05-1020 draft (80 frags)	6441.43	6441.43	2718	1848
Enterobacteriaceae	Yersinia pestis	Yersinia pestis biovar Orientalis str. MG05-1020 draft (80 frags)	6441.43	6441.43	2718	1848

Figure 3.2 Detection results with two identical target sequences

This could occur as a result of:

- an author submitting the identical sequence to a given repository on multiple occasions
- an author submitting the identical sequence to multiple different public domain repositories
- the incorporation of sequences from one public domain repository into another

The automated KPATH pipeline is unable to determine that these sequences are identical in these cases.

From the example in Figure 3.2 we check the database for source information about the two detected target sequences and it would appear that the two target sequences represent the same sequence that were downloaded from NCBI on two different dates under two different source identifiers (Table 3.1):

Table 3.1 Identical target sequences present in MiDAS HybDB as two different target sequences.

kp|730838	kp|2518984|730838|Glued fragments of sequence 730739 (Yersinia pestis biovar Orientalis str. MG05-1020 Yersinia pestis biovar Orientalis str. MG05-1020, whole genome shotgun sequencing project from NCBI on May 07 2010 01:26A...) - 80 fragments[sequence_id 730838][seq_data_id 2518984][tax_node_id 404217] Glued fragments of sequence 730739 (Yersinia pestis biovar Orientalis str. MG05-1020 Yersinia pestis biovar Orientalis str. MG05-1020, whole genome shotgun sequencing project from NCBI on May 07 2010 01:26A...) - 80 fragments
kp|847926	kp|9379705|847926|Glued fragments of sequence 843288 (Yersinia pestis biovar Orientalis str. MG05-1020 Yersinia pestis biovar Orientalis str. MG05-1020, whole genome shotgun sequencing project from NCBI on Nov 07 2013 2:45PM) - 80 fragments[sequence_id 847926][seq_data_id 9379705][tax_node_id 404217] Glued fragments of sequence 843288 (Yersinia pestis biovar Orientalis str. MG05-1020 Yersinia pestis biovar Orientalis str. MG05-1020, whole genome shotgun sequencing project from NCBI on Nov 07 2013 2:45PM) - 80 fragments

Draft genome sequences

Target sequences with "draft" designations are often whole genome shotgun sequencing database submissions that contain multiple contigs. The Axiom MiDAS algorithm will occasionally give preference to these "draft" target sequences over finished sequences when making detection calls. The contigs in these "draft" records may cover the majority of multiple chromosomes or plasmids for a particular organism. As a consequence these "draft" records will often have more associated probes and may thus attain a higher conditional score compared to any of the corresponding "finished" target sequences representing only a single chromosome or plasmid.

In the Figure 3.3, a draft target sequence for *B. thailandensis* E264 was detected. This draft sequence (kp|849989) likely contains contigs that cover both chromosomes of *B. thailandensis* E264 (3.8 Mb and 2.9 Mb, respectively, http://www.straininfo.net/genomes/10774).

Family	Species	Target Description	Conditional Score	Initial Score	Probes Expected	Probes Observed
Burkholderiaceae	**Burkholderia thailandensis**	**Burkholderia thailandensis E264 draft (44 frags)**	**2049.42**	**2049.42**	**837**	**578**
Burkholderiaceae	**Burkholderia thailandensis**	**Burkholderia thailandensis E264 strain ATCC 700388 draft (44 frags)**	**2049.42**	**2049.42**	**837**	**578**
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis Bt4 draft (803 frags)	2032.95	2032.95	887	585
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis Bt4 draft (803 frags)	2032.18	2032.18	888	585
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E555 draft (521 frags)	1954.56	1954.56	870	566
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis TXDOH draft (810 frags)	1941.72	1941.72	861	561
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis TXDOH draft (810 frags)	1941.67	1941.67	861	561
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E264 chromosome I	1350.27	1350.27	559	381

Figure 3.3 Draft sequences detected

The detected draft sequence (kp849998) likely covers the majority of both chromosomes. There are 837 corresponding probes for it, and 578 of which were observed as exceeding the threshold intensities (Figure 3.4):

Family	Species	Target Description	Conditional Score	Initial Score	Probes Expected	Probes Observed
Burkholderiaceae	**Burkholderia thailandensis**	**Burkholderia thailandensis E264 draft (44 frags)**	**2049.42**	**2049.42**	**837**	**578**
Burkholderiaceae	**Burkholderia thailandensis**	**Burkholderia thailandensis E264 strain ATCC 700388 draft (44 frags)**	**2049.42**	**2049.42**	**837**	**578**
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis Bt4 draft (803 frags)	2032.95	2032.95	887	585
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis Bt4 draft (803 frags)	2032.18	2032.18	888	585
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E555 draft (521 frags)	1954.56	1954.56	870	566
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis TXDOH draft (810 frags)	1941.72	1941.72	861	561
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis TXDOH draft (810 frags)	1941.67	1941.67	861	561
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E264 chromosome I	1350.27	1350.27	559	381

Figure 3.4 Expected and observed probes for primary hit

For the highest scoring 'secondary' hit to a finished sequence for *B. thailandensis* E264 chromosome I, we notice the number of probes expected is approximately half that for the detected draft target sequence (Figure 3.5).

Family	Species	Target Description	Conditional Score	Initial Score	Probes Expected	Probes Observed
Burkholderiaceae	**Burkholderia thailandensis**	**Burkholderia thailandensis E264 draft (44 frags)**	**2049.42**	**2049.42**	**837**	**578**
Burkholderiaceae	**Burkholderia thailandensis**	**Burkholderia thailandensis E264 strain ATCC 700388 draft (44 frags)**	**2049.42**	**2049.42**	**837**	**578**
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis Bt4 draft (803 frags)	2032.95	2032.95	887	585
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis Bt4 draft (803 frags)	2032.18	2032.18	888	585
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E555 draft (521 frags)	1954.56	1954.56	870	566
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis TXDOH draft (810 frags)	1941.72	1941.72	861	561
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis TXDOH draft (810 frags)	1941.67	1941.67	861	561
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E264 chromosome I	1350.27	1350.27	559	381

Figure 3.5 Expected and observed probes for secondary hit, chromosome I.

Likewise, for the first 'secondary' hit to *B. thailandensis* E264 chromosome II, the number of expected probes is also approximately half that for the detected draft sequence (Figure 3.6).

Family	Species	Target Description	Conditional Score	Initial Score	Probes Expected	Probes Observed
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E264 chromosome II	1056.32	1056.32	441	299
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis 2007721723 chromosome 2	1046.6	1046.6	440	297

Figure 3.6 Expected and observed probes for secondary hit, chromosome II.

Note: In the process of designing probes against the sequences in the KPATH database, the individual contigs in each of the "draft" sequences were joined together with 'N' characters to ensure that no probes were designed against chimeras.

Brackets around taxonomic names

In some instances the detected target sequences may have taxonomic names (family, genus, species) enclosed in square brackets (e.g., *[Clostridium] stercorarium*). The square brackets represent recommended taxonomies (mainly based on genome trees), but not verified taxonomies
(see: https://groups.google.com/forum/#!topic/qiime-forum/WnhQiB5q9Hc)
(Figure 3.7).

Family	Species	Target Description
Lachnospiraceae	[Clostridium] saccharolyticum	Clostridium cf. saccharolyticum K10
Eubacteriaceae	[Eubacterium] eligens	Eubacterium eligens ATCC 27750 plasmid unnamed
Lachnospiraceae	[Ruminococcus] torques	Ruminococcus torques L2-14 draft genome

Figure 3.7 Taxonomic names with brackets

Algorithm details

Calculating the initial composite likelihood

The initial composite likelihood for a target J is:

$$\sum_{i:Yi=1} \log\left(\frac{P(Y_i = 1 \mid X_j = 1)}{P(Y_i = 1 \mid X_j = 0)}\right) + \sum_{i:Yi=0} \log\left(\frac{P(Y_i = 0 \mid X_j = 1)}{P(Y_i = 0 \mid X_j = 0)}\right)$$

Where:

- Y_i is the i^{th} probe, and X_j is the j^{th} Target.
- $Y_i = 1$ indicates the probe is above threshold
- $X_j = 1$ indicates the target is present
- And $P(Y \mid X)$ is the probability that the probe Y is in a state given that the sample contains or does not contain target X

Calculating the conditional likelihood

When the conditional likelihood for each subsequent target is calculated, it is calculated with the assumption that the prior targets are already in the sample, so the conditional likelihood now reads:

$$\sum_{i:Y_{i=1}} log\left(\frac{P\left(Y_i = 1 \middle| X_j = 1, X_k = 1 \forall k \in S\right)}{P\left(Y_i = 1 \middle| X_j = 0, X_k = 1 \forall k \in S\right)}\right)$$
$$+ \sum_{i:Y_{i=0}} log\left(\frac{P(Y_i = 0 | X_j = 1, X_k = 1 \forall k \in S)}{P(Y_i = 0 | X_j = 0, X_k = 1 \forall k \in S)}\right)$$

Where:

- Y_i is the i^{th} probe, and X_j is the j^{th} Target.
- $Y_i = 1$ indicates the probe is above threshold
- $X_j = 1$ indicates the target is present
- And $P(Y|X)$ is the probability that the probe Y is in a state given that the sample contains or does not contain target X
- k is an individual prior target that was detected in an earlier iteration
- S is the set of all prior targets that were detected in earlier iteration

Once we assume that all of the prior targets are in the sample, $P(Y_i = 1 | X_j = 0)$ no longer has a value of about .01, but now has a value >0.5. In other words, each individual probe that already has a reason to be bright can't give much support to the hypothesis that a new target is in the sample.

3 Best practices workflow

Grouping Samples

In contrast to Axiom™ genotyping products, there is no lower limit on the number of samples to be analyzed at a given time to produce high quality data. Sample analysis is self-contained using both probes interrogating microbes as well as background control probes on a single Axiom™ Microbiome Array. However, reasonable and good experimental study design (controls, appropriate replicates, etc.) are recommended.

It is also recommended to group together samples that will be compared to each other.

Perform analysis on process control sample

In addition to providing a process control for the Axiom™ 2.0 Target Preparation, the included REF103 human gDNA sample is also used to monitor the fidelity of array processing and array performance. In addition to the probes designed for microbial interrogation, Axiom Microbiome Array also includes probes targeted towards regions of the human genome, and are used to generate a metric termed Dish QC (DQC).

DQC is based on intensities of probe sequences for non-polymorphic human genome locations (i.e., sites that do not vary in sequence from one individual to the next). When subject to the two-color Axiom assay, probes expected to ligate an A or T base (referred to as AT non-polymorphic probes) produce specific signal in the AT channel and background signal in the GC channel. The converse is true for probes expected to ligate a G or C base (referred to as GC non-polymorphic probes). DQC is a measure of the resolution of the distributions of "contrast" values, where:

$$\text{Contrast} \sim = \frac{AT\ Signal - GC\ Signal}{AT\ Signal + GC\ Signal}$$

Distributions of contrast values are computed separately for the AT non-polymorphic probes (which should produce positive contrast values) and GC non-polymorphic probes (which should produce negative contrast values). If the assay was successfully executed, then signal will be high in the expected channel and low in background channel, and the two contrast distributions will be well-resolved. A DQC value of zero indicates no resolution between the distributions of AT and GC probe contrast values, and the value of 1 indicates perfect resolution. A DishQC value for REF103 less than 0.82 indicates a probable issue with general performance of all samples on the plate.

4 Analysis of Axiom™ MiDAS results

Summary table

The summary table view allows an initial analysis of the sample composition with regards to both the number of targets detected by the Axiom™ Microbial Detection Analysis Software (MiDAS), as well how these targets can be summarized at the three most discrete levels of taxonomic resolution (Figure 3.8). The taxonomic identification is generated from the Axiom MiDAS target descriptions to give an unweighted diversity, or count of the number of Families, Genera, and Species, detected in the sample. This view allows the cross-comparison of the samples, which can be further visualized using the Family-level tiles.

Sample Name	Array Type	Number of Targets Detected	Number of Species Detected	Number of Genera Detected	Number of Families Detected
NTC.CEL	Axiom_MDA	0	0	0	0
REF103.CEL	Axiom_MDA	0	0	0	0
43589D2_Axiom_MDA_D01.CEL	Axiom_MDA	2	2	2	2
43589D2_Axiom_MDA_D09.CEL	Axiom_MDA	2	2	2	2
700388D5_Axiom_MDA_B01.CEL	Axiom_MDA	3	3	2	2
700388D5_Axiom_MDA_B09.CEL	Axiom_MDA	3	2	2	2
5spMix_Axiom_MDA_F03.CEL	Axiom_MDA	10	7	5	5
5spMix_Axiom_MDA_F11.CEL	Axiom_MDA	9	6	5	5
HM276D_Axiom_MDA_H03.CEL	Axiom_MDA	41	30	17	18
HM276D_Axiom_MDA_H04.CEL	Axiom_MDA	38	28	17	18
Stool1_Axiom_MDA_A03.CEL	Axiom_MDA	52	44	32	18
Stool2_Axiom_MDA_B03.CEL	Axiom_MDA	50	41	29	18
NR43976_Axiom_MDA_B05.CEL	Axiom_MDA	4	2	2	1
NR43976_Axiom_MDA_D07.CEL	Axiom_MDA	3	2	2	1
VR106_Axiom_MDA_B07.CEL	Axiom_MDA	4	3	2	2
VR106_Axiom_MDA_D05.CEL	Axiom_MDA	4	3	2	2

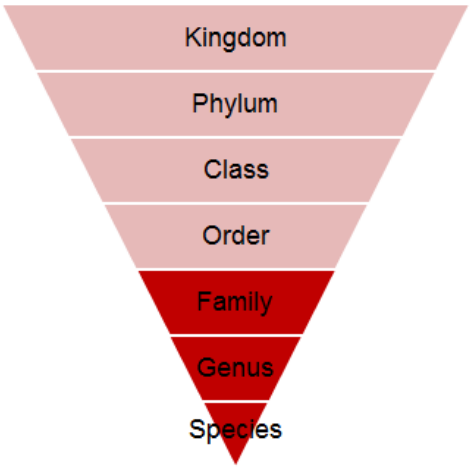


Figure 3.8 Summary table

In addition to providing the initial view of experimental samples, the summary view table also allows the quick inspection of a no template control (NTC) as well as the supplied process control, REF103. Any targets detected in a NTC sample may be microbial contaminants, and thus these targets may not be present in other samples in which they are detected.

Family-level visualization

In the **Family** tab, the Family-level tiles allow visualization of Axiom™ Microbiome Array data anywhere from one, to all of the samples that have been interrogated by Axiom MiDAS. Each tile represents one of 560 families represented on the array, which are in a fixed position in the grid. If none of the samples are selected from the summary table, the default is for the family-level tiles for all samples to be rendered. Initial inspection of this view can be used to make a high-level visual comparison of the sample (Figure 3.9).

Sample Name	Array Type	Number of Targets Detected	Number of Species Detected	Number of Genera Detected	Number of Families Detected
NTC.CEL	Axiom_MDA	0	0	0	0
REF103.CEL	Axiom_MDA	0	0	0	0
43589D2_Axiom_MDA_D01.CEL	Axiom_MDA	2	2	2	2
43589D2_Axiom_MDA_D09.CEL	Axiom_MDA	2	2	2	2
700388D5_Axiom_MDA_B01.CEL	Axiom_MDA	3	3	2	2
700388D5_Axiom_MDA_B09.CEL	Axiom_MDA	3	2	2	2
5spMix_Axiom_MDA_F03.CEL	Axiom_MDA	10	7	5	5
5spMix_Axiom_MDA_F11.CEL	Axiom_MDA	9	6	5	5
HM276D_Axiom_MDA_H03.CEL	Axiom_MDA	41	30	17	18
HM276D_Axiom_MDA_H04.CEL	Axiom_MDA	38	28	17	18
Stool1_Axiom_MDA_A03.CEL	Axiom_MDA	52	44	32	18
Stool2_Axiom_MDA_B03.CEL	Axiom_MDA	50	41	29	18
NR43976_Axiom_MDA_B05.CEL	Axiom_MDA	4	2	2	1
NR43976_Axiom_MDA_D07.CEL	Axiom_MDA	3	2	2	1
VR106_Axiom_MDA_B07.CEL	Axiom_MDA	4	3	2	2
VR106_Axiom_MDA_D05.CEL	Axiom_MDA	4	3	2	2

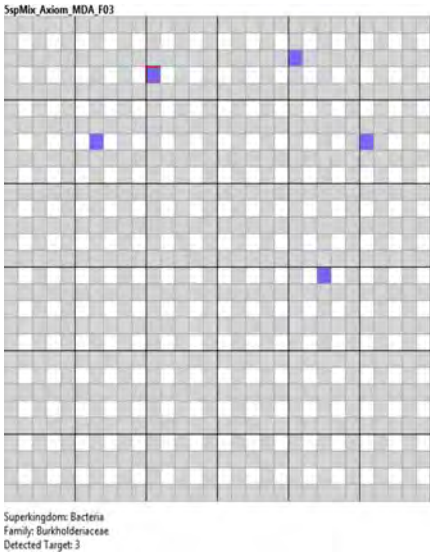


Figure 3.9 Family-level visualization

When a family is detected in a sample, the corresponding tile is colored in blue. The family-level tile is indicated regardless of the number of targets that are included in the family. To determine the family, the tile can be selected, indicated by a red outline around the border of the tile. This selects the corresponding family-level tile for all of the samples selected in the summary table. The tile selected yields three pieces of information, the Superkingdom and Family of the tile, and the number of targets that have been detected within that Family.

Condensed family-level visualization

The **Condensed Family** tab gives an alternative view of the Family-level description of samples interrogated with Axiom MiDAS. In this view, only tiles that represent families detected by Axiom MiDAS in one or more of the samples analyzed are rendered. The view is organized such that family tiles are organized by the superkingdom to which they belong in alphabetical order, which are delineated by different rows in the viewer. Detected families within the row are organized in alphabetical order with each row representing a different superkingdom (Figure 3.10). As in the **Family** tab, one to all of the samples can be selected from the sample summary table to be viewed. As before, selecting one of the tiles yields that same information as can be found on the family tiles.

Sample Name	Array Type	Number of Targets Detected	Number of Species Detected	Number of Genera Detected	Number of Families Detected
NTC.CEL	Axiom_MDA	0	0	0	0
REF103.CEL	Axiom_MDA	0	0	0	0
43589D2_Axiom_MDA_D01.CEL	Axiom_MDA	2	2	2	2
43589D2_Axiom_MDA_D09.CEL	Axiom_MDA	2	2	2	2
700388D5_Axiom_MDA_B01.CEL	Axiom_MDA	3	3	2	2
700388D5_Axiom_MDA_B09.CEL	Axiom_MDA	3	2	2	2
5spMix_Axiom_MDA_F03.CEL	Axiom_MDA	10	7	5	5
5spMix_Axiom_MDA_F11.CEL	Axiom_MDA	9	6	5	5
HM276D_Axiom_MDA_H03.CEL	Axiom_MDA	41	30	17	18
HM276D_Axiom_MDA_H04.CEL	Axiom_MDA	38	28	17	18
Stool1_Axiom_MDA_A03.CEL	Axiom_MDA	52	44	32	18
Stool2_Axiom_MDA_B03.CEL	Axiom_MDA	50	41	29	18
NR43976_Axiom_MDA_B05.CEL	Axiom_MDA	4	2	2	1
NR43976_Axiom_MDA_D07.CEL	Axiom_MDA	3	2	2	1
VR106_Axiom_MDA_B07.CEL	Axiom_MDA	4	3	2	2
VR106_Axiom_MDA_D05.CEL	Axiom_MDA	4	3	2	2

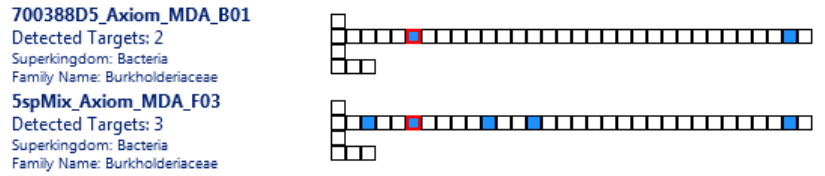


Figure 3.10 Condensed Family-level Visualization

Microbe frequency view

The microbe frequency view depicts the frequency of microbes appearing in the sample set at two different levels of taxonomic resolution. On the x-axis the user can view either the families or species detected on at least one sample interrogated by the Axiom™ Microbiome Array. The y-axis gives the percentage of the families or species in the samples. The user can also determine which sample the family/species of interest is detected in by clicking its corresponding bar. This highlights the sample(s) containing the family/species of interest (Figure 3.11). Note that when REF103 and a no template control are run, no family or species should reach a frequency of 100%, as these two samples should be devoid of any microbes.

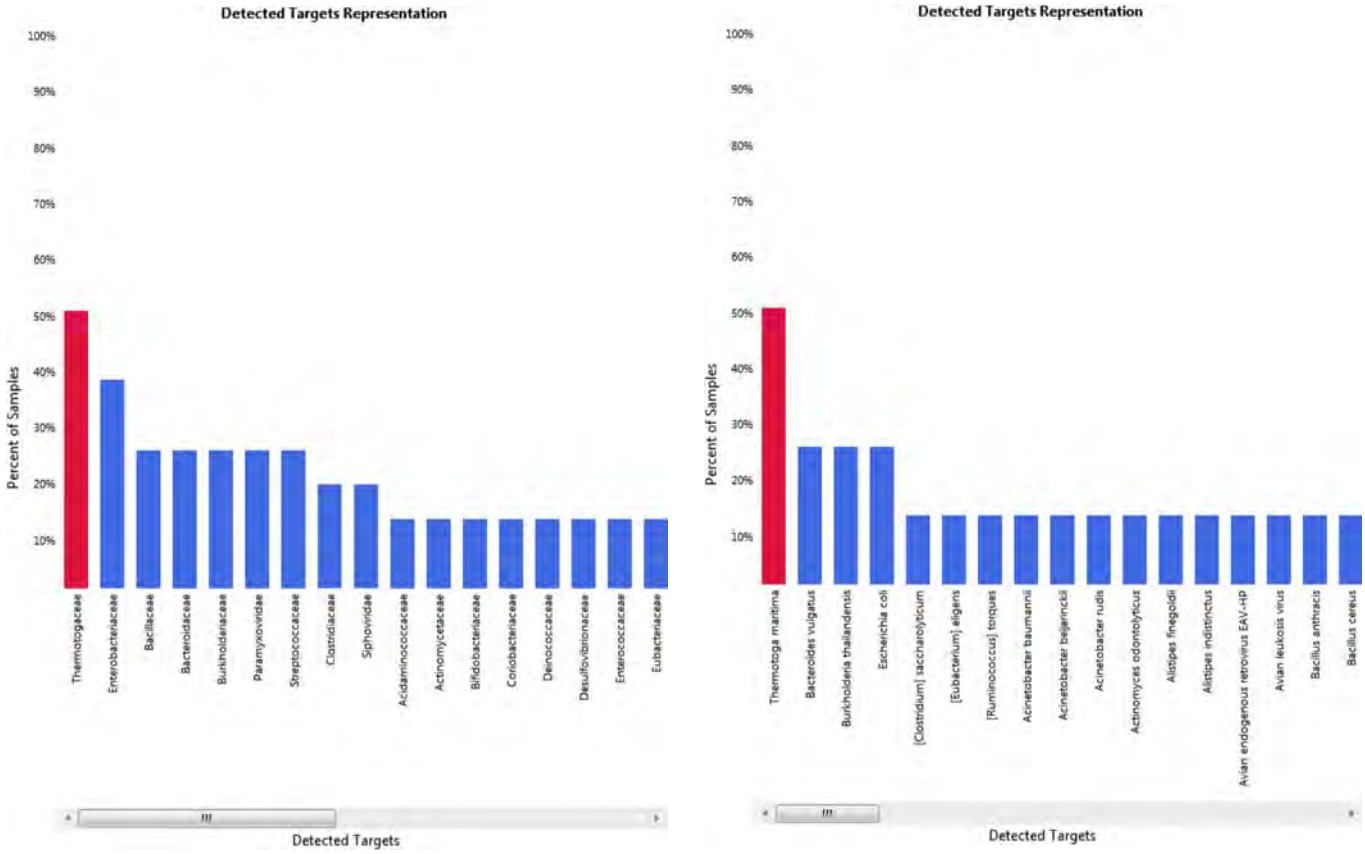


Figure 3.11 Microbe Frequency View

Sample view

From the **Summary** tab the user can select a specific sample of interest and use the **Sample** tab to investigate the targets that have been identified by Axiom MiDAS to yield the best predicted model of the probe-level data. The views that are rendered are the lists of the targets that make up the sample, a graph of the log of conditional scores vs. the ratio of conditional to initial scores. The family tile initially viewed in the **Summary** tab is rendered below the graphical output.

Sample summary list

The default columns provide the summary data of each target over three levels of taxonomic hierarchy. The user can use the Show/Hide columns tab to yield data on the other taxonomic levels (e.g., Phylum) on the table (Figure 3.12).

Family	Species	Target Description	Conditional Score	Initial Score	Probes Expected	Probes Observed
Enterobacteriacea	Yersinia pestis	Yersinia pestis biovar Orientalis str. MG05-1020 draft (8	6441.43	6441.43	2718	1848
Bacillaceae	Bacillus anthracis	Bacillus anthracis str. 95014 draft (27 frags)	2731.48	2731.48	1386	839
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis Bt4 draft (803 frags)	2055.18	2055.18	887	590
Francisellaceae	Francisella philomiragia	Francisella philomiragia strain FAJ draft (6 frags)	431.58	442.93	181	127
Thermotogaceae	Thermotoga maritima	Thermotoga maritima MSB8	165.22	165.22	108	56
Enterobacteriacea	Yersinia pestis	Yersinia pestis D106004 plasmid pPCY1	8.37	36.13	19	11
Burkholderiaceae	Burkholderia sp. JPY347	Burkholderia sp. JPY347 draft (60 frags)	7.72	235.58	206	92
Francisellaceae	Francisella philomiragia	Francisella philomiragia subsp. philomiragia ATCC 25017	4.26	173.14	63	48
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E264 chromosome II	1.33	1101.88	441	308
Francisellaceae	Francisella noatunensis	Francisella noatunensis subsp. orientalis LADL--07-285A	0.82	212.07	100	63

Figure 3.12 Sample summary list

The default list of the target data has eight columns providing information about each target. The complete list of column table definitions is found in Table 3.2.

Table 3.2

Column Name	Description
Target ID	Unique identifier for the target.
Superkingdom	Superkingdom classification of the detected target.
Phylum	Phylum classification of the detected target.
Class	Class classification of the detected target.
Order	Order classification of the detected target.
Family	Family classification of the detected target.
Species	Species classification of the detected target.
Target Description	Database record of the detected target.
Conditional Score	The conditional likelihood ratio for a target if detected targets from previous iterations are present.
Iteration	The analysis round in which the target was chosen.
Initial Score	The likelihood ratio for a target if no other targets are present.
Probes Expected	Number of distinct probes expected to give a positive signal for the detected target.
Probes Observed	Number of distinct probes giving a positive signal for the detected target.
Total Expected Probe Hyb Probability	Sum of the hybridization (detection) probabilities of distinct expected probes.
Total Observed Probe Hyb Probability	Sum of the hybridization (detection) probabilities of distinct observed probes.

The Target Description of the sample is a list of targets from the database that best explain the observed probe signals. Each target is the best database match of the probes that yielded this hit and may provide more discreet information about the target (e.g., strain). The user can also use the **Link To** button to learn more information about any specific target from NCBI. This feature provides a link to at least one of the following NCBI databases:

- **Bioproject**—Project driven database of biological data derived from a single organization or consortium.
- **GenBank**—NIH database containing all publicly available annotated genetic sequences.
- **Taxonomy**—Database of organism names and phylogenetic relationships.

Figure 3.13 Linking to Outside Databases

In our example above "Thermotoga maritima MSB" is selected. If the user uses the link to the Genbank sequence, this opens the browser to the matching record at NCBI (Figure 3.13).

Following the taxonomic identities, the default view provides the scores and number of supporting probes used to determine the sample's Target Description. The distinct probe counts have unique target genomic coordinates and refer to those used in the conditional and initial score calculations.

Initial and conditional scores

As described in the introduction of this guide, the initial score is the likelihood ratio that the target is present in the sample if no other targets are present in the sample. This value gives information on what the *maximum possible contribution* of that target is to the holistic model of the sample, based on the probes observed when interrogating a sample with the Axiom™ Microbiome Array. The conditional score gives a value of the *actual contribution* of each target to the model of the sample. As Axiom MiDAS takes into account the presence of other targets, this can modify the conditional score of a given target if there are probes in common between targets.

Detected and secondary targets: Using the conditional score percentile slider

As Axiom MiDAS selects the target that best explains probe-level data (Detected Target), there can be targets that have conditional scores that are quite close and were not selected (Secondary Targets). These Secondary Targets are within the same taxonomic Family as the Detected Target. The description of the Secondary Target can be assessed from the target list in the **Sample** tab using the Conditional Score Percentile Slider. This tool allows the viewing of Secondary Targets within the percentile of the Detected Target within the iteration of Axiom MiDAS in which it is detected, either by using the slider or by input of a number into Conditional Score percentile (set at 100% by default) (Figure 3.14).

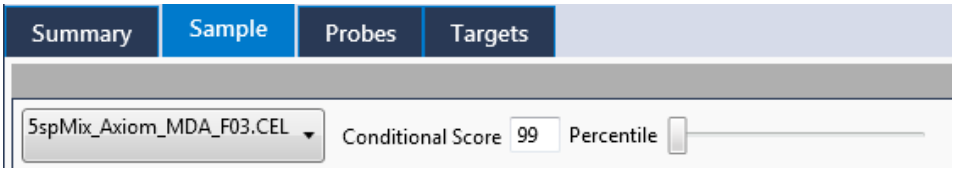


Figure 3.14 Visualize secondary targets: Conditional Score percentile slider

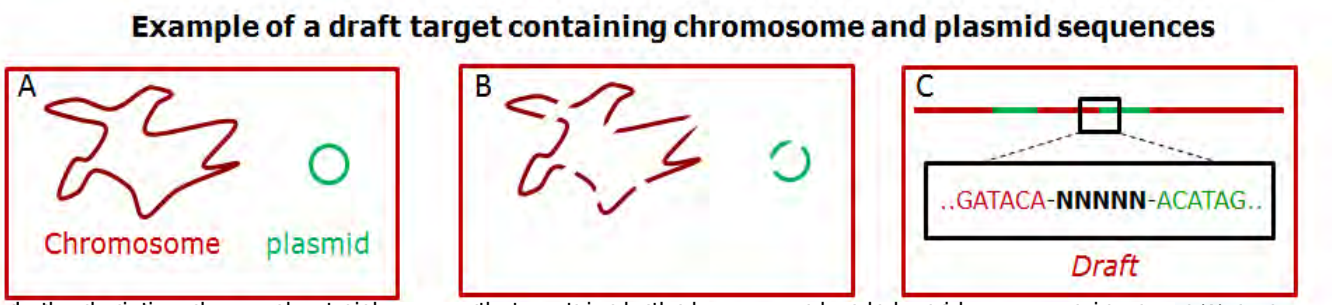
This allows the user to determine if there were Secondary Targets in the iteration and how close their conditional score was to the Detected Target. For example if the conditional slider is set to 99% of a sample that contains *Burkholderia thailandensis*, strain E264, the Detected Target that was selected by Axiom MiDAS is *Burkholderia thailandensis*, strain Bt4, but when we investigate the Secondary Targets we note that the E264 draft has a conditional score that is only slightly less than the Detected Target (Figure 3.15). This example illustrates that Secondary Targets with conditional scores that are only slightly lower than the scores for the Detected Targets should be evaluated as possible alternatives for the Detected Targets. This is especially true in the case of draft sequences described in the next section.

Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis Bt4 draft (803 frags)	2055.18	2055.18
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis Bt4 draft (803 frags)	2054.41	2054.41
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E264 draft (44 frags)	2051.23	2051.23
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E264 strain ATCC 700388 draft (44 frags)	2051.23	2051.23

Figure 3.15 Visualize Secondary Targets: Sample Table

Targets containing draft sequences

When viewing the Target Descriptions in the Sample View, note that Detected or Secondary Target descriptions can contain descriptors from draft genomes. These are denoted by the terminology "draft" in the description followed by a parenthesis and the number of fragments used to generate that draft genome. In the Axiom MiDAS target database, fragments from draft genomes have been "glued" together with an intervening number of "N's". These intervening N's were placed in the sequence to ensure that probes were not designed over these glued boundaries (Figure 3.16).



In the depiction above, a bacterial genome that contains both chromosomal and plasmid sequences is depicted (A). In the case of draft target, fragments of sequence (B) have been assembled together with a run of intervening N's (C) in the HybDB database.

Figure 3.16 Example of a draft target containing chromosome and plasmid sequences.

As bacterial genomes can have multiple chromosomal and plasmid sequences, it is important to note that draft genomes can harbor information from these sequences. The implication of this is that draft genomes can have higher conditional score than curated genomes where multiple chromosomes and plasmids are separated into different target sequences. If the Detected Target is a draft sequence, it is recommended to evaluate the Secondary Targets, as there may be a non-draft target sequence with a conditional score that is only slightly lower than the draft sequence. In the scenario where the draft sequence conditional score is inflated due to the presence of more probes, the non-draft target sequence in the Secondary Targets may be of the actual present strain.

Using the Sample Graph: Log of conditional scores vs. ratio of conditional/initial scores

In the default sample view the target-level data is rendered in a graph depicting the Log of Conditional Scores vs. the Ratio of Conditional/Initial Scores. The intended use of this graphical output is to allow the user to analyze the relative contribution of the targets to the model of the data provided by Axiom MiDAS. We will first describe analysis of this tool in the default view, where only Detected Targets are depicted (Figure 3.17), and then discuss how using the Conditional Score Slider changes the view of the data.

Family	Species	Target Description	Conditional Score	Initial Score	Probes Expected	Probes Observed
Enterobacteriacea	Yersinia pestis	Yersinia pestis biovar Orientalis str. MG05-1020 draft (E	6441.43	6441.43	2718	1848
Bacillaceae	Bacillus anthracis	Bacillus anthracis str. 95014 draft (27 frags)	2731.48	2731.48	1386	839
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis Bt4 draft (803 frags)	2055.18	2055.18	887	590
Francisellaceae	Francisella philomiragia	Francisella philomiragia strain FAJ draft (6 frags)	431.58	442.93	181	127
Thermotogaceae	Thermotoga maritima	Thermotoga maritima MSB8	165.22	165.22	108	56
Enterobacteriacea	Yersinia pestis	Yersinia pestis D106004 plasmid pPCY1	8.37	36.13	19	11
Burkholderiaceae	Burkholderia sp. JPY347	Burkholderia sp. JPY347 draft (60 frags)	7.72	235.58	206	92
Francisellaceae	Francisella philomiragia	Francisella philomiragia subsp. philomiragia ATCC 25017	4.26	173.14	63	48
Burkholderiaceae	Burkholderia thailandensis	Burkholderia thailandensis E264 chromosome II	1.33	1101.88	441	308
Francisellaceae	Francisella noatunensis	Francisella noatunensis subsp. orientalis LADL--07-285A	0.82	212.07	100	63

Figure 3.17 Default Sample View, 0.2 Ratio of Conditional/Initial Scores.

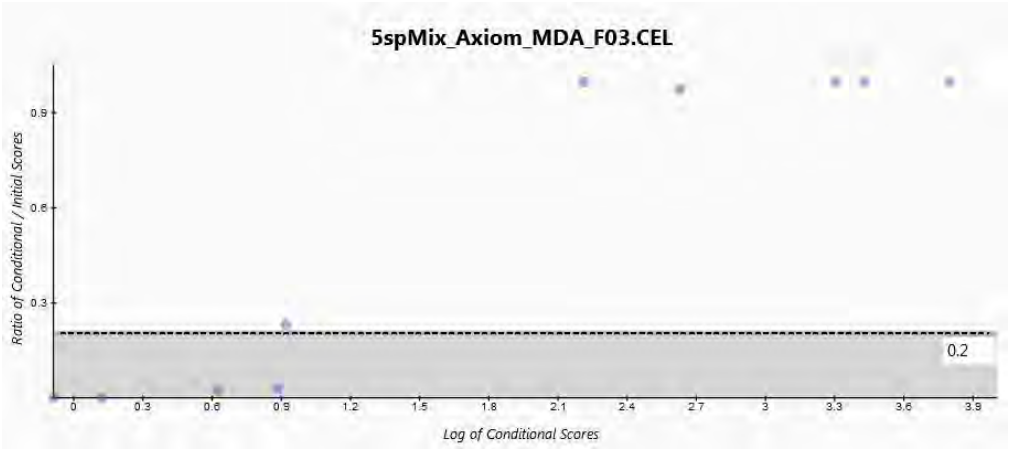


Figure 3.18 Default Sample View Graph, 0.2 Ratio of Conditional/Initial Scores.

The Log of Conditional Scores vs. Ratio of Conditional/Initial Scores graph allows the user to examine target level data in two respects (Figure 3.18). The x-axis (Log of Conditional Scores) provides information on the *relative contribution of each target* to the model provided by Axiom MiDAS, while the y-axis provides the ratio of the actual probe-level data used to contribute to the model relative to its possible maximal contribution. The actual contribution of a target sequence to the model may be lower than its possible maximal contribution because a similar target sequence, with shared probes, has already been detected and therefore explained the shared probe signals.

Using the Plot Graph Threshold tool

On the sample viewer graph, the Plot Graph Threshold tool is provided to the user. This threshold can be set in one of two ways, either by dragging the line up from the x-axis, or by entering a value between 0-1 into the box at the right corner of the graph. As this value is increased, targets that have a Log Ratio of Conditional/Initial Score below the threshold change from black to gray on the sample table. This tool allows users to view targets and their relative and actual contribution to the Axiom MiDAS generated data model for the sample.

A threshold of 0.2 is the default for the Ratio of Conditional/Initial Scores. This setting suggests that targets having 80% (or more) of their initial log likelihood score explained by prior hits on the chip are unlikely to be present as distinct targets in the sample. The exception to this recommendation is target sequences for plasmids. These plasmid targets should be considered detected even with low ratios values, as these low values may be due to probe overlap with detected targets for their host bacteria.

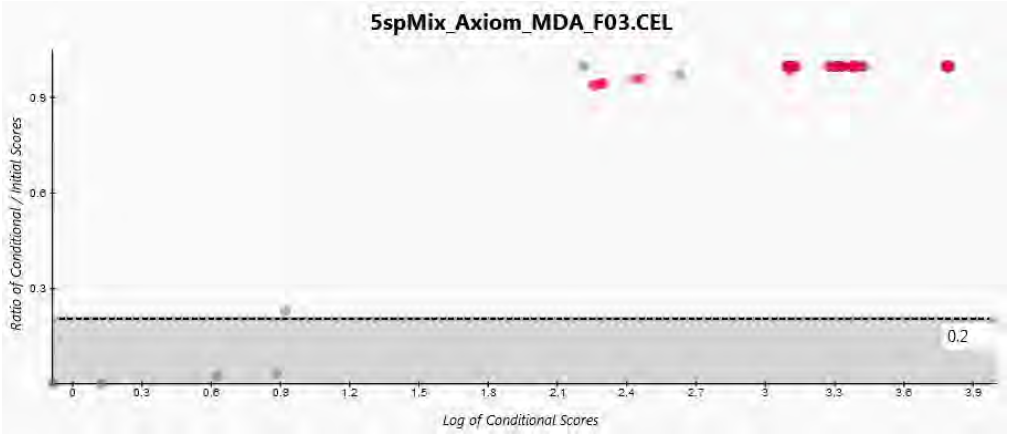


Figure 3.19 Conditional Score Slider Tool

When using the Conditional Score Slider Tool, Secondary Targets also populate this graph (see Figure 3.19). These can be distinguished from Detected Targets as they are in red rather than light blue.

Probes View

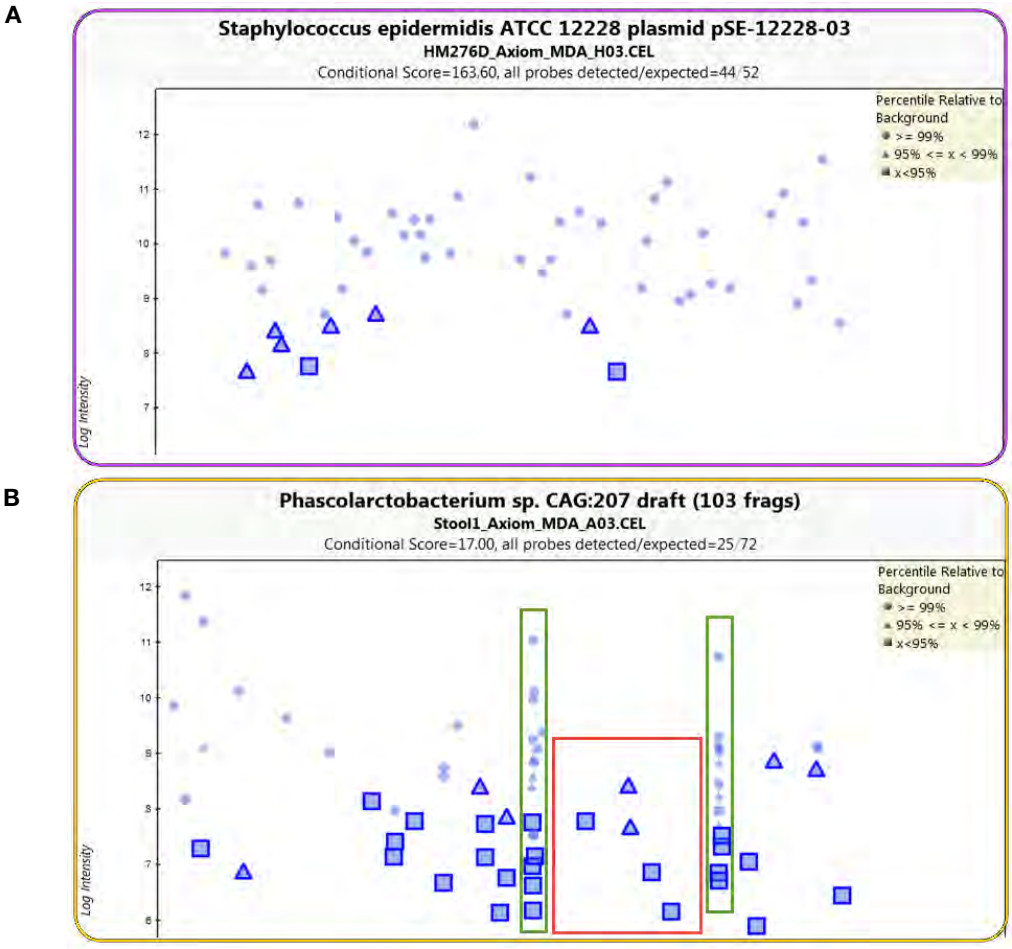


Figure 3.20 Probes View

The **Probes** window tab (Figure 3.20) displays probes for detected targets from a sample selected in the Sample view. Its graph provides data to support the Targets Detected by the Axiom MiDAS application. This graph is plotted as the log intensity of probes detected for a given target versus the position of the probe on the target sequence. This view is can be informative to determine if probe responses are even over the length of a target sequence, or if there are certain regions that have poor performance. If there are regions that are missing along the length of a target sequence, this could indicate the presence of a sequence variant.

In the default analysis mode, only the probes that are above the 99th percentile (circles) contribute positively to the log likelihood score of the target. Probes that are below the 99th percentile (triangles and squares) subtract from the log likelihood of the target.

The outlined probes in Figure 3.20 (triangles and squares) are below the 99th percentile of background probes and, therefore, contribute negatively to the highlighted Conditional Score for each respective target. (A) Good probe signal coverage across target and few probes below the 99th percentile. (B) Probe signal concentrated on specific regions, possible indication of sequence variant, conserved region outlined by green boxes or under-performing regions outlined by red box.

Targets View

Target Description	NTC	REF103	43589D2_ Axiom_M DA_D01	43589D2_ Axiom_M DA_D09	700388D5 _Axiom_ MDA_B01	700388D5 _Axiom_ MDA_B09	5spMix_A xiom_MD A_F03	5spMix_A xiom_MD A_F11
Pyrococcus sp. NA2 chromosome			Secondary	Secondary				
Pyrococcus furiosus COM1 chromosome			DETECTED	DETECTED				
Pyrococcus horikoshii OT3 chromosome			Secondary	Secondary				
Pyrococcus furiosus DSM 3638 chromosome			Secondary	Secondary				
Thermotoga sp. RQ2 chromosome			Secondary	Secondary	Secondary	Secondary	Secondary	Secondary
Thermotoga sp. EMP draft (13 frags)			Secondary	Secondary	Secondary	Secondary	Secondary	Secondary
Thermotoga petrophila RKU-1 chromosome			Secondary	Secondary	Secondary	Secondary	Secondary	Secondary
Thermotoga maritima MSB8 draft (7 frags)			DETECTED	Secondary	Secondary	Secondary	Secondary	Secondary
Thermotoga neapolitana DSM 4359 chromosom			Secondary	Secondary	Secondary	Secondary	Secondary	Secondary
Thermotoga naphthophila RKU-10			Secondary	Secondary	Secondary	Secondary	Secondary	Secondary
Thermotoga maritima MSB8			Secondary	DETECTED	DETECTED	DETECTED	DETECTED	DETECTED

Figure 3.21 Targets View

The **Targets Window** tab displays target calls across all your analyzed samples. Its table format allows you to quickly search for targets of interest and export your data for use in other applications. Primary and Secondary Targets are identified by "DETECTED" and "Secondary" in the table respectively (Figure 3.21).

Safety

⚠ **WARNING! GENERAL SAFETY**. Using this product in a manner not specified in the user documentation may result in personal injury or damage to the instrument or device. Ensure that anyone using this product has received instructions in general safety practices for laboratories and the safety information provided in this document.

- Before using an instrument or device, read and understand the safety information provided in the user documentation provided by the manufacturer of the instrument or device.
- Before handling chemicals, read and understand all applicable Safety Data Sheets (SDSs) and use appropriate personal protective equipment (gloves, gowns, eye protection, etc). To obtain SDSs, see the "Documentation and Support" section in this document.

Chemical safety

⚠ **WARNING!** GENERAL CHEMICAL HANDLING. To minimize hazards, ensure laboratory personnel read and practice the general safety guidelines for chemical usage, storage, and waste provided below, and consult the relevant SDS for specific precautions and instructions:

- Read and understand the Safety Data Sheets (SDSs) provided by the chemical manufacturer before you store, handle, or work with any chemicals or hazardous materials. To obtain SDSs, see the "Documentation and support" section in this document.
- Minimize contact with chemicals. Wear appropriate personal protective equipment when handling chemicals (for example, safety glasses, gloves, or protective clothing).
- Minimize the inhalation of chemicals. Do not leave chemical containers open. Use only with adequate ventilation (for example, fume hood).
- Check regularly for chemical leaks or spills. If a leak or spill occurs, follow the manufacturer's cleanup procedures as recommended in the SDS.
- Handle chemical wastes in a fume hood.
- Ensure use of primary and secondary waste containers. (A primary waste container holds the immediate waste. A secondary container contains spills or leaks from the primary container. Both containers must be compatible with the waste material and meet federal, state, and local requirements for container storage.)
- After emptying a waste container, seal it with the cap provided.
- Characterize (by analysis if necessary) the waste generated by the particular applications, reagents, and substrates used in your laboratory.
- Ensure that the waste is stored, transferred, transported, and disposed of according to all local, state/provincial, and/or national regulations.
- **IMPORTANT!** Radioactive or biohazardous materials may require special handling, and disposal limitations may apply.

Biological hazard safety

⚠ **WARNING!** **BIOHAZARD.** Biological samples such as tissues, body fluids, infectious agents, and blood of humans and other animals have the potential to transmit infectious diseases. All work should be conducted in properly equipped facilities using the appropriate safety equipment (for example, physical containment devices). Safety equipment also may include items for personal protection, such as gloves, coats, gowns, shoe covers, boots, respirators, face shields, safety glasses, or goggles. Individuals should be trained according to applicable regulatory and company/ institution requirements before working with potentially biohazardous materials. Follow all applicable local, state/ provincial, and/or national regulations. The following references provide general guidelines when handling biological samples in laboratory environment.

- U.S. Department of Health and Human Services, *Biosafety in Microbiological and Biomedical Laboratories (BMBL)*, 5th Edition, HHS Publication No. (CDC) 21-1112, Revised December 2009; found at:
  www.cdc.gov/biosafety/publications/bmbl5/BMBL.pdf
- World Health Organization, *Laboratory Biosafety Manual*, 3rd Edition, WHO/ CDS/CSR/LYO/2004.11; found at:
  www.who.int/csr/resources/publications/biosafety/Biosafety7.pdf

Documentation and support

Related documentation

Document	Publication number
Axiom 2.0 Assay 24-Array Format Manual Workflow	
Axiom™ 2.0 Assay 24-Array Format Manual Workflow User Guide	703335
Axiom™ 2.0 Assay 24-Array Format Manual Workflow Site Preparation Guide	703336
Axiom™ 2.0 Assay 24-Array Format Manual Workflow Quick Reference	703337
Axiom 2.0 Assay 96-Array Format Manual Workflow	
Axiom™ 2.0 Assay 96-Array Format Manual Workflow User Guide	702990
Axiom™ 2.0 Assay 96-Array Format Manual Workflow Site Preparation Guide	702991
Axiom™ 2.0 Assay 96-Array Format Manual Workflow Quick Reference	702989
Axiom 2.0 Assay 96-Array Format Automated Workflow—NIMBUS	
Axiom™ 2.0 Assay 96-Array Format Automated Workflow User Guide—Applied Biosystems NIMBUS™ Instrument	MAN0017740
Axiom™ 2.0 Assay 96-Array Format Automated Workflow Site Preparation Guide—Applied Biosystems NIMBUS™ Instrument	703350
Axiom™ 2.0 Assay 96-Array Format Automated Workflow Quick Reference—Applied Biosystems NIMBUS™ Instrument	703351
Supporting documentation	
Axiom™ 2.0 gDNA Sample Prep Protocol Quick Reference	702987
GeneTitan™ MC Protocol for Axiom™ 2.0 Array Plate Processing Quick Reference	702988
GeneTitan™ Multi-Channel Instrument User Guide	08-0308
GeneTitan™ Multi-Channel Instrument Site Preparation Guide	08-0305
GeneChip™ Command Console Software User Guide	702569

Customer and technical support

Visit **thermofisher.com/support** for the latest in services and support, including:

- Worldwide contact telephone numbers
- Product support, including:
  - Product FAQs
  - Software, patches, and updates
- Order and web support
- Product documentation, including:
  - User guides, manuals, and protocols
  - Certificates of Analysis
  - Safety Data Sheets (SDSs; also known as MSDSs)

    Note: For SDSs for reagents and chemicals from other manufacturers, contact the manufacturer.

Limited product warranty

Life Technologies Corporation and/or its affiliate(s) warrant their products as set forth in the Life Technologies' General Terms and Conditions of Sale found on Life Technologies' website at **www.thermofisher.com/us/en/home/global/terms-and-conditions.html**. If you have any questions, please contact Life Technologies at **thermofisher.com/support**.