isospeak®

User Guide

Legal Notices

Revision History

Document #704-00016-01

Rev. 1.12

IsoSpeak Ver 3.0.0

Table of Contents

Chapter 1: Overview

Introduction

IsoSpeak is an integrated data analysis software package for the IsoLight and IsoSpark instrument platforms. It provides a comprehensive and user-friendly suite of bioinformatics tools for the downstream analysis of raw IsoLight/IsoSpark image data. Explore changes in protein expression and secretion in each sample, identify trends and differences between groups, and easily visualize your results with IsoSpeak.

Features

Project Annotation	Easy-to-use panels to record information about each sample/experiment in a project
Data Processing	Automated data processing and quality control to convert raw IsoLight/IsoSpark image data into accurate, numerical data
Data Analysis	Integrated analysis software processes to analyze, visualize, and stratify a project's data

Additional Resources

The following documentation is available for download and viewing from the Bruker website:

Resource	Description
IsoSpeak Overview	An overview of the IsoSpeak software, including interactive videos.
Analyze the Data	Analyzing proteomic data using IsoSpeak.

IsoSpeak Terminology

Terminology	Description
IsoSpeak Project	A saved file with a .spkpro extension and associated subfolder containing all the data and annotations for a project. This may consist of data from one or more runs on the IsoLight/IsoSpark.
IsoSpeak Computer	A separate computer for initial experiment planning and for secondary data analysis using IsoSpeak once a run of the IsoLight/IsoSpark has been completed.
Experiment	An experiment corresponds to a single chip analyzed on the IsoLight/IsoSpark.
Run	A run consists of up to 8 experiments run in one batch on the IsoLight/IsoSpark.
IsoLight/IsoSpark Raw Data	Image data output by the IsoLight/IsoSpark after completion of a run, which is imported into IsoSpeak for subsequent annotation, data processing and analysis.
Project	A project consists of a set of related experiments (can be more than one run).
Project Info	Annotation interface in IsoSpeak to record information about all the experiments and samples in the project.
Data Processing	Experiment data from the IsoLight/IsoSpark is automatically processed during this stage, converting image data into quantifiable, highly multiplexed single-cell or bulk readouts.
Data Analysis	Bioinformatics tools and visualizations for analysis and exporting of the data from one or more experiments or an entire project.

System Requirements

Description	Recommended Specification
Operating System	Windows 10 or Windows 11
Processor	Intel® Core™ i3 2.5 GHz or similar (i5 or i7 recommended)
Graphics Card	NVIDIA® GeForce® GT 620M, ATI Radeon™ HD 4670 or Intel® HD Graphics 4400 or similar
Memory	8 GB RAM (16 GB recommended)
Storage	500 GB available hard drive space (1 TB recommended)
Screen Resolution	1920 x 1080 display resolution (recommended)

Installing IsoSpeak

IsoSpeak will come preinstalled on the IsoSpeak Computer.

If you need to install or reinstall IsoSpeak to your computer, please log in to Citrix ShareFile and go to the IsoSpeak Updates folder (see Chapter 13 for more information).

To install, download the .msi file to the computer and double click the .msi installation file, and follow the installation instructions. If the installer prompts you to allow the app to make changes to your device, please select “yes”.

Chapter 2: Using IsoSpeak's Informatics Pipeline for Single-Cell Data Analysis

Differentiate and stratify using single-cell polyfunctional indices.

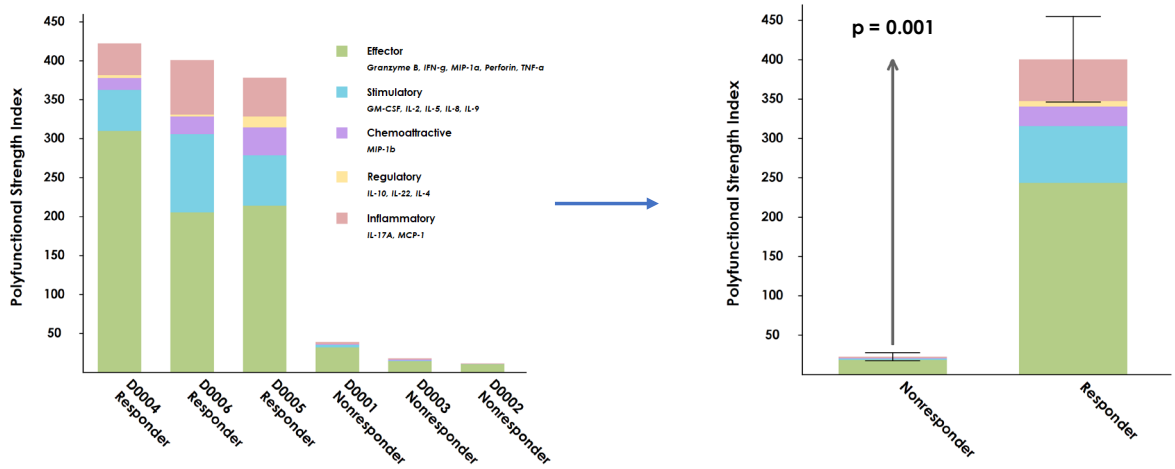
IsoSpeak gives you the power to use polyfunctional indices such as the Polyfunctional Index (PSI/PII), which aggregate all single-cell, multi-dimensional secretions from a sample into a single index and provides you an overall quality metric for each of your samples.

The readout combines the polyfunctionality of a sample (frequency of cells secreting multiple cytokines) with the signal intensities for each single cell across the secreted cytokines of the sample. The displayed index is color-coded to show the contribution from different categories of cytokines (e.g., effector vs. stimulatory cytokines).

This metric can be used for comparing the relative response of a set of samples or sample groups, such as

1. Samples from patients responding to a therapy vs. non-responders.
2. Samples produced using manufacturing process A vs. B
3. Samples from patients treated with combination immunotherapy A vs. B
4. Samples from patients with neuroinflammatory progression vs. control patients

Polyfunctional indices can be viewed individually per heterogeneous sample or aggregated over sample groups to shed light on significant differences between groups.



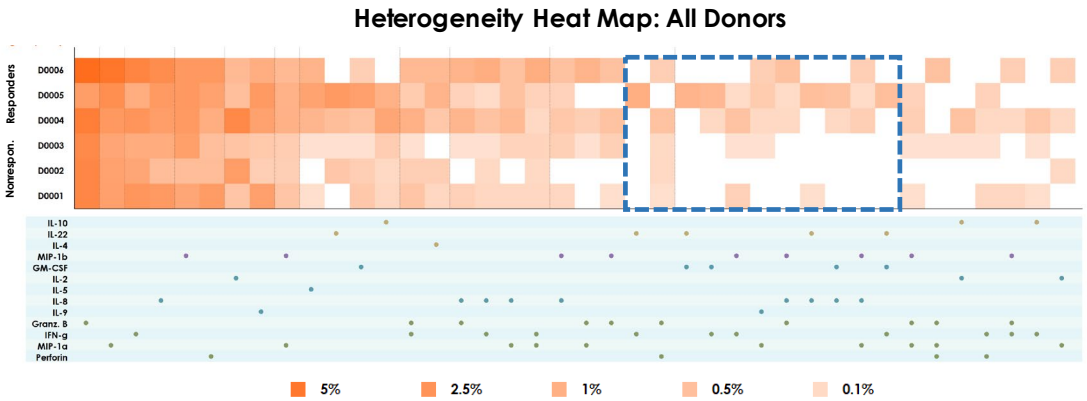
IsoSpeak's Project Info module allows you to easily annotate each patient sample. These annotations automatically translate to selection filters that you can use to aggregate samples and compare across sample groups, guiding the entire single-cell informatics pipeline to tease out sample group differences.

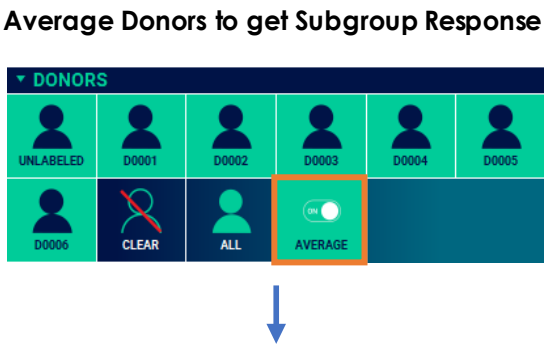
Read more in Chapter 9.

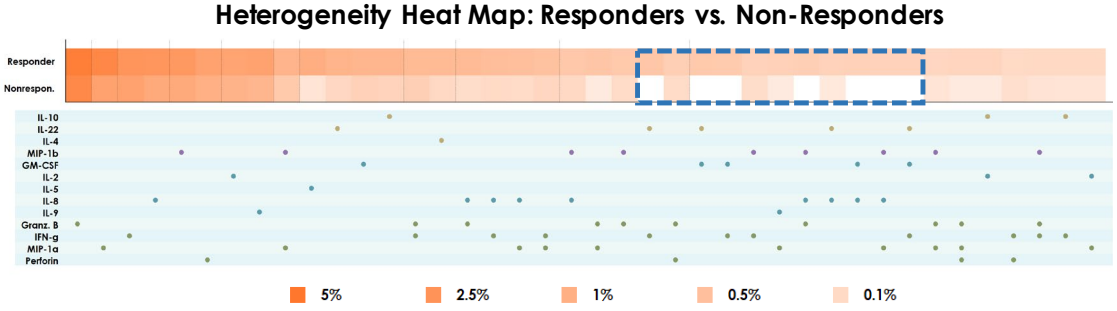
Stratify subpopulations that are driving differences.

When differences are seen across samples groups, a critical step is then uncovering the drivers of these differences using IsoSpeak. You can uncover, process, replicate and utilize the underlying single-cell subsets that drive differences in patient cohorts.

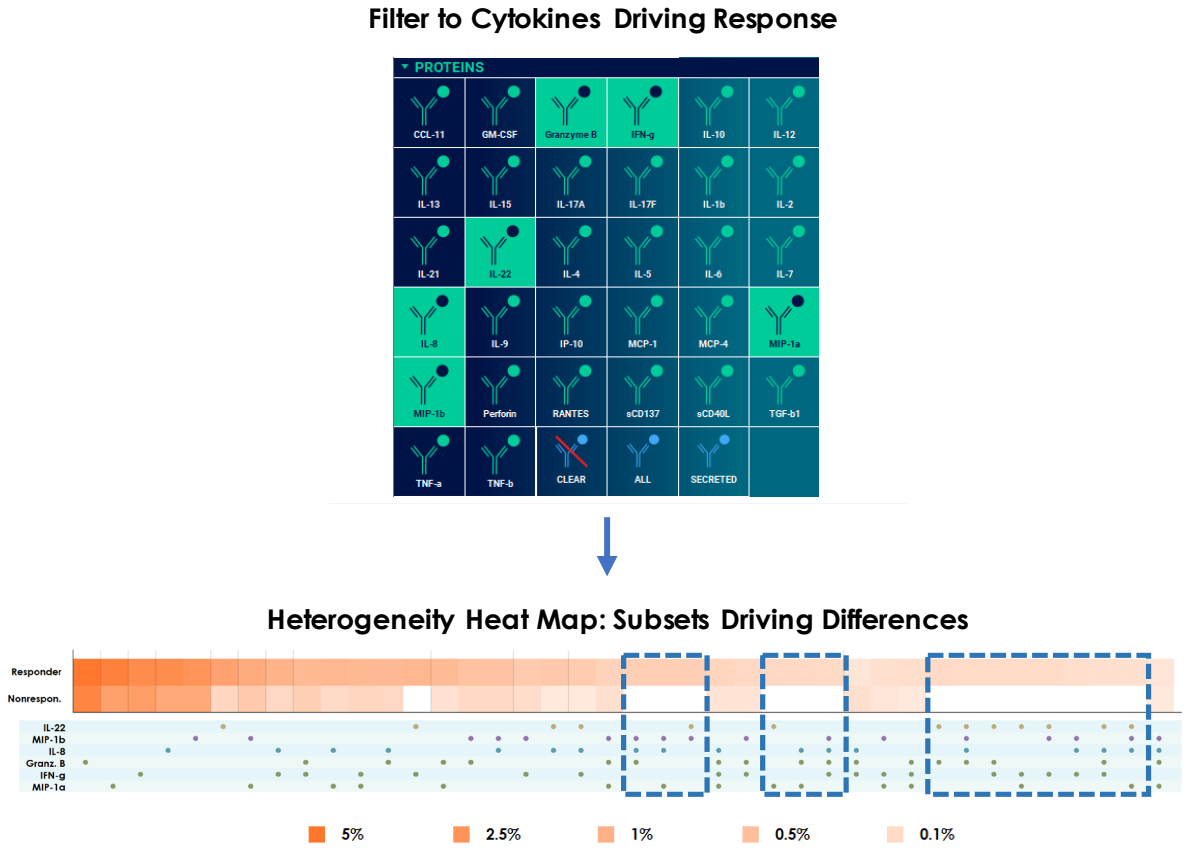
a. Use Heterogeneity Heat Maps to uncover the critical cell subpopulations that exist only in the condition/group of interest.







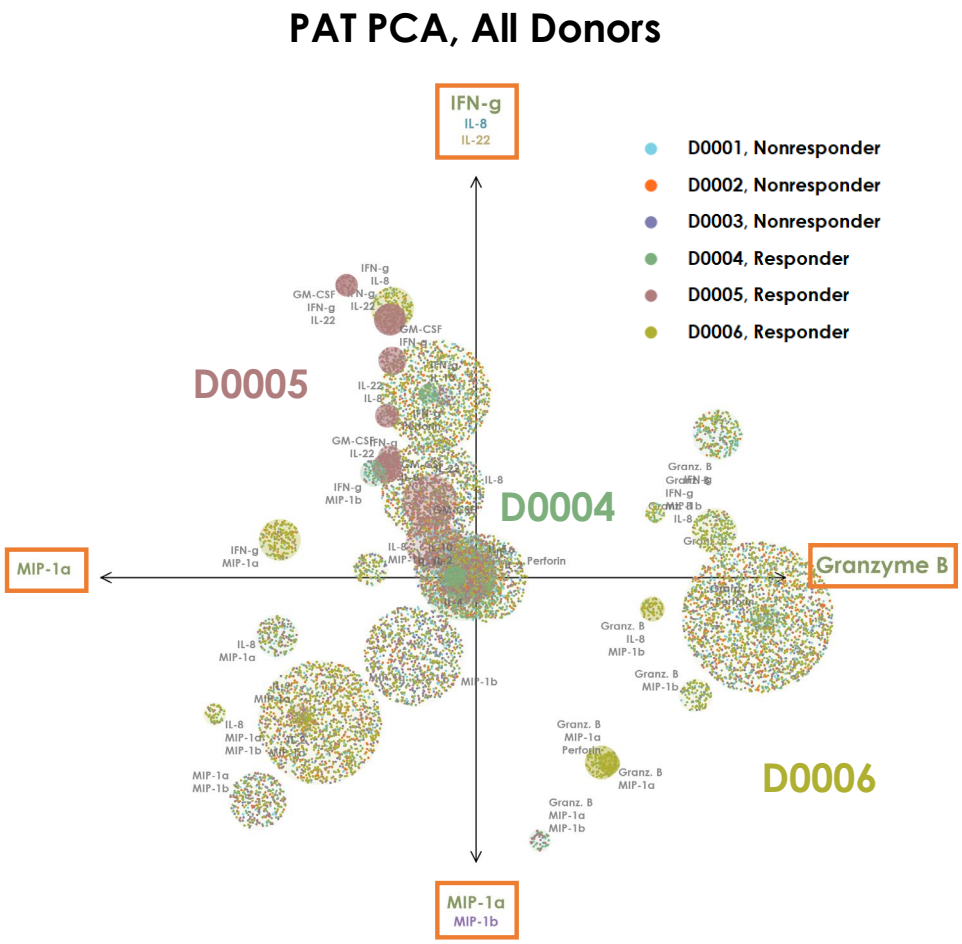
Then, narrow down the analyzed cytokines to zero in on the multi-dimensional subsets that tell the story.



Read more about heat maps in Chapter 9.

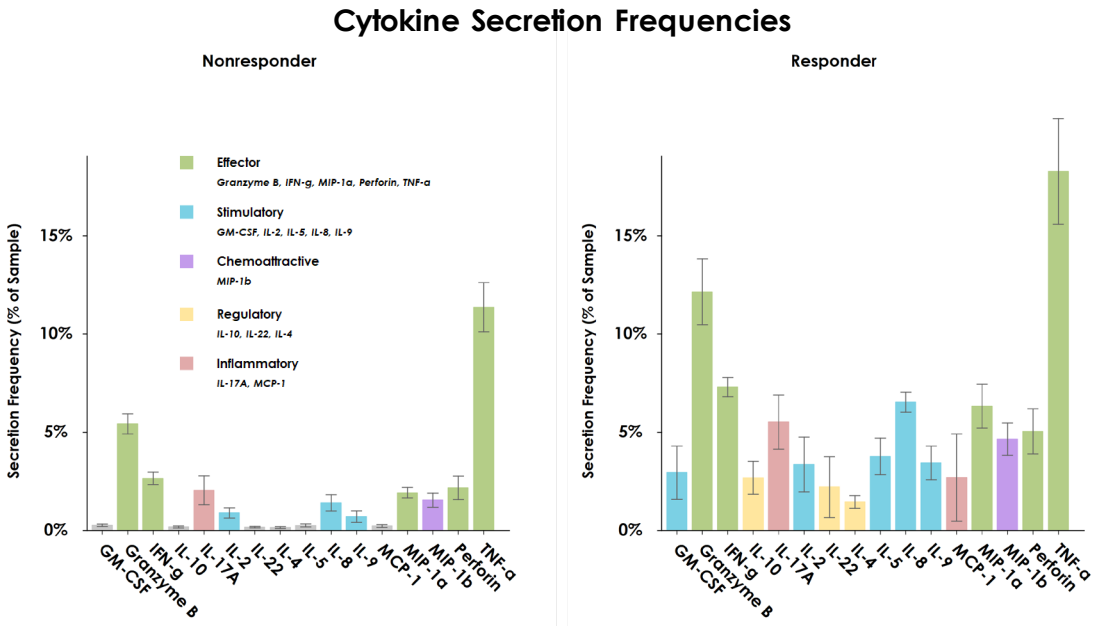
Recognize subgroups to enable multi-functional cytokine stratification.

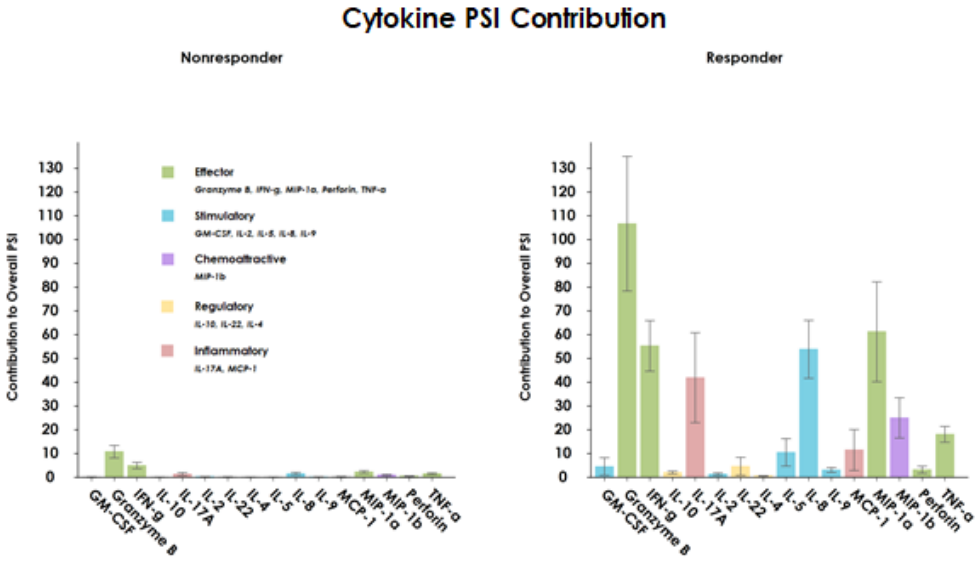
Dominant subgroups will emerge in PAT PCA graphs, representing significant multi-functional subsets driving the overall response. Samples with higher response dominate the graph, in this case all corresponding to the "responder" subgroup. Differences between these three donors are highlighted through the location of their polyfunctional subsets in the graph.



Read more about PAT PCA in Chapter 9.

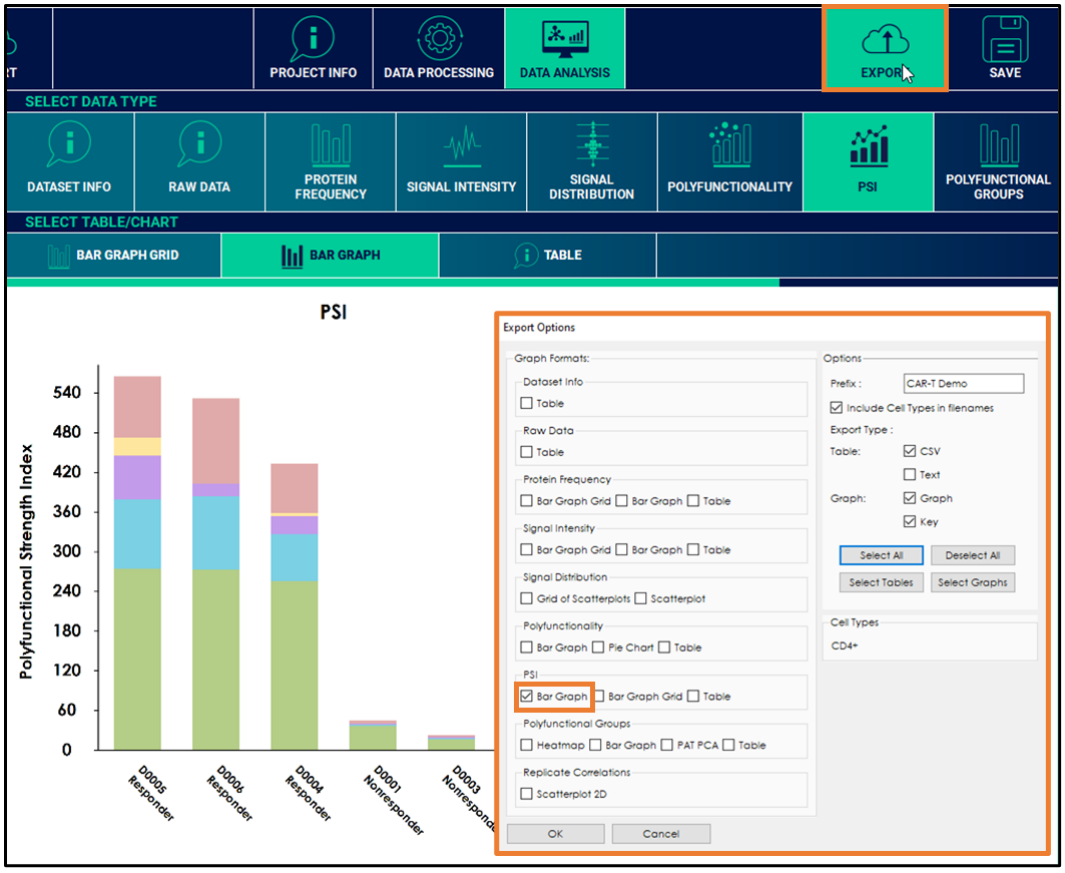
Verify the largest components of those subgroups and look holistically at your data.





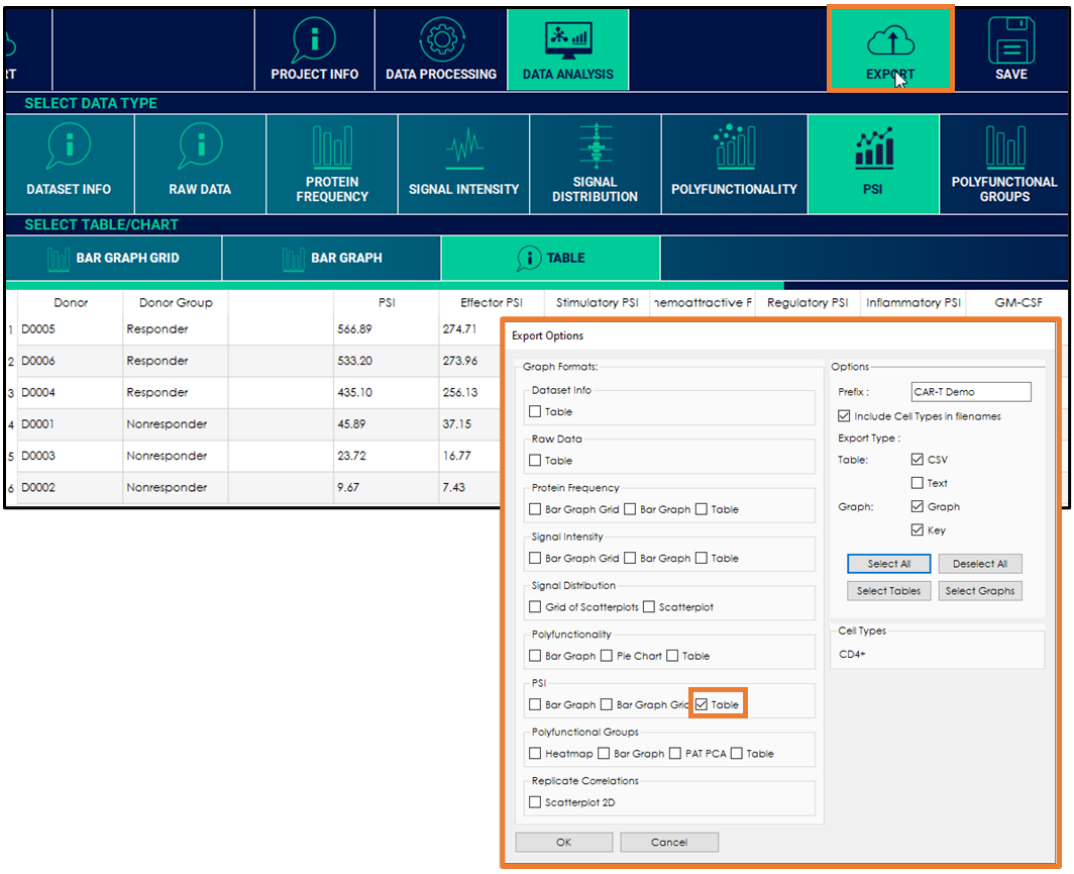
Read more about these graphs in Chapter 9.

Export visualized datasets from your pipeline.



For more details on data exporting, see Chapter 8.

Add to your pipeline with your own analytics.



For more details on data exporting, see Chapter 8.

Starting IsoSpeak

You should be able to find a shortcut to IsoSpeak on your desktop or under the start menu. Look for the following icon:

and double-click to start. You should see the following screen pop up:

New Project

Start filling in information for a new project and set of experiments you plan to or have run on the IsoLight/IsoSpark.

Load Project

Load an existing project (.spkpro project file).

Recent Projects

Click to load a recently accessed project.

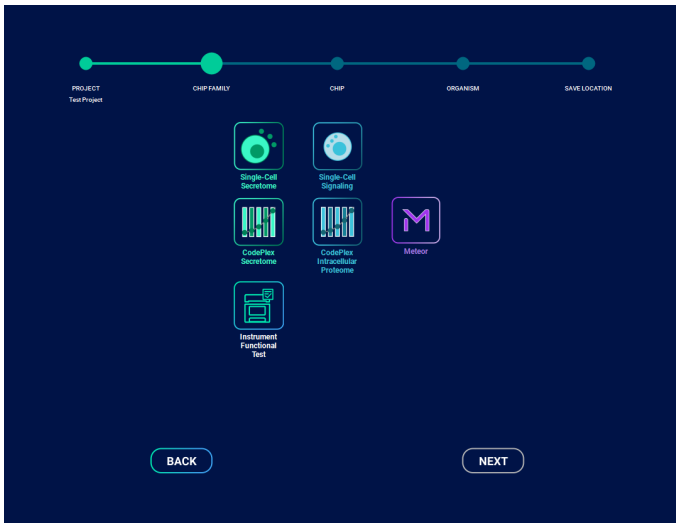
Starting a Project

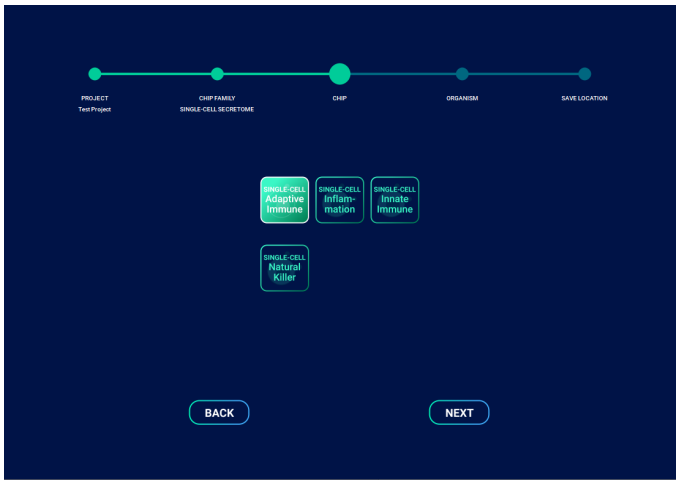
To start a new project, click the New Project button. Enter the name of the project and click on the next button. For the three successive screens, select the following:

- Chip Family (e.g., Single-Cell Secretome/IsoCode),
- Chip (corresponding to the panel of proteins being analyzed, e.g., Adaptive Immune), and
- Organism (e.g., Human) on successive screens.

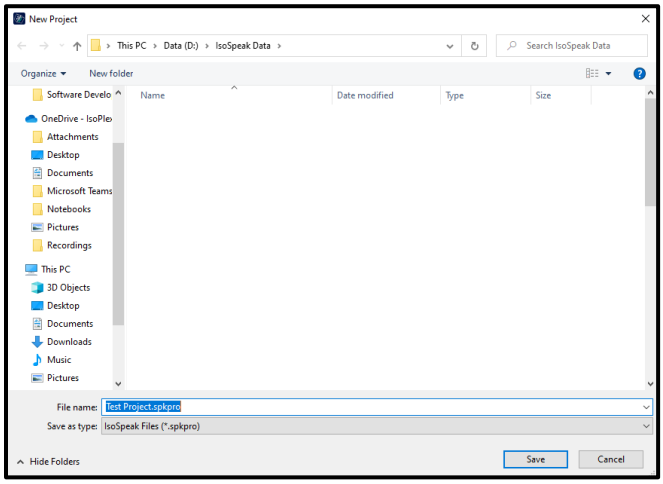
An IsoSpeak project can only contain data from a specific application (chip type), i.e., all data and IsoLight/IsoSpark runs imported into this project must correspond to this application.

Use “Next” to move to the next selection. At any time, you can go “Back” to change your selection.

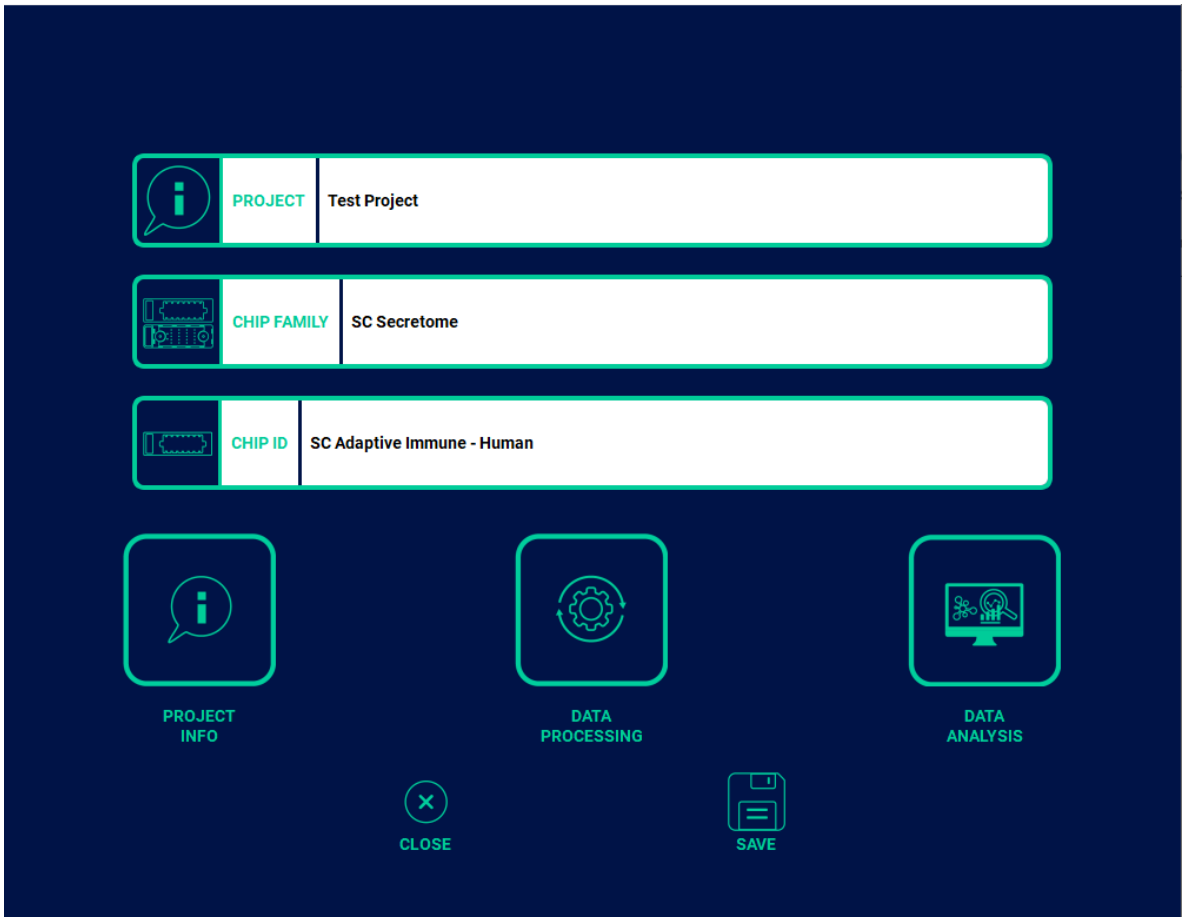




After selecting Organism, you will see a browser window for you to specify the name and choose the location of your project. The default location of all IsoSpeak projects is under **\Documents\IsoSpeak Data**. This location can be changed by going to **Options, Set IsoSpeak Data Location…** and choosing a different folder.



Once a project is created, you will see the Main Menu, also shown below. The three buttons in the middle correspond to the three stages of the project analysis workflow.



The next step is to start recording the experiment(s) that are associated with this project. This can be done in the Project Info analysis stage, described in more detail on the following pages.

Clicking the Save button will save the project and its experiments.

Clicking the Close button exits out of the project.

Project Analysis Workflow

There are three general stages in the IsoSpeak workflow:

The Project Info module allows you to import raw IsoLight or IsoSpark data into your project and annotate each of these corresponding samples.

The Data Processing module automatically translates the raw image data from the IsoLight/IsoSpark into numerical data.

The Data Analysis module contains tables and visualizations of the resulting numerical data from your project. This module provides tools to comprehensively analyze and explore the data in the project, as well as compare and stratify various subsets of the data.

Chapter 4: Importing IsoCode Single-Cell Raw Data and Project Annotation

Under Project Info, each experiment corresponds to one chip run on the IsoLight or IsoSpark. Once the raw image data is acquired from an instrument run, you will need to import, annotate, and process this data through IsoSpeak to obtain quantified data (e.g., single-cell, cytokine secretion data). The first two steps, outlined in this chapter, involve importing raw data into the IsoSpeak project and then annotating your experiments.

Importing IsoCode Single-Cell Raw Data

To add a new set of experiments that were run on the IsoLight/IsoSpark, go to the Project Info module, and click the Import button at the top of the screen.

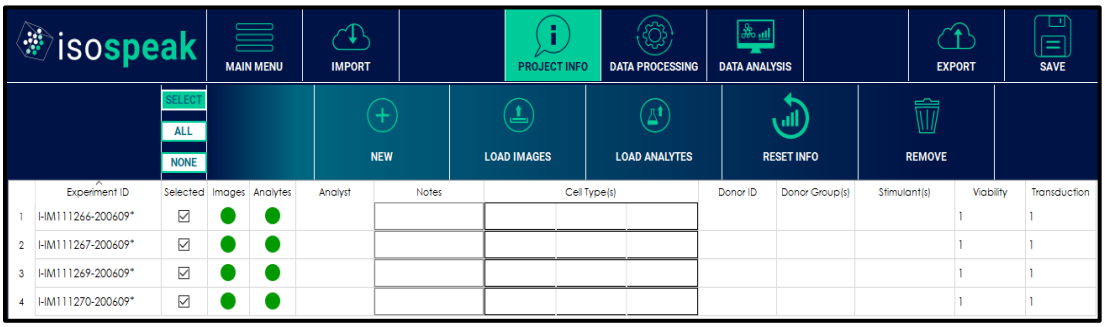
In the subsequent popup, select Instrument Data.

Use the browser to select the folder containing the raw IsoLight/IsoSpark data you wish to import and click OK.

The raw data will then get copied to the project's file structure, under the project's Images folder (see Chapter 13 IsoSpeak Project File Structure for details).

Note: this step might take a while (for 4 chips expect ~15 min when importing from a USB 3.0 drive)

Once imported, the new experiments should automatically appear in the table, with their raw images and analyte information already uploaded.



The import status of the raw data is reflected by the circular Images and Analytes buttons of each experiment (under the Images or Analytes column in the table).

The import status colors are as follows for the buttons:

Green: the raw data (images or analytes) was loaded successfully.

Blue: the raw data (images or analytes) has not yet been input or has not been found.

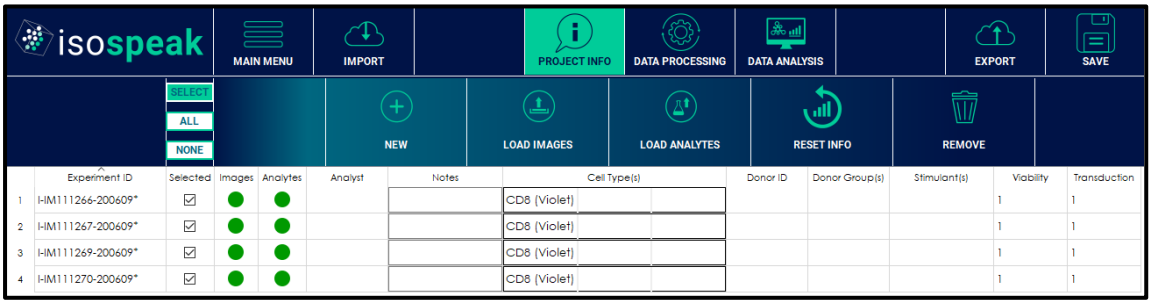
Red: the previously input raw data (images or analytes) cannot be located anymore.

If necessary, raw image data and/or analyte information can be loaded individually after the experiments are created. The data must already reside under the project's Images folder for this to succeed (see Chapter 13 IsoSpeak Project File Structure for details).

To load an individual experiment's images or analytes, click the experiment's circular Images or Analytes button. To batch-load images or analytes for multiple experiments, select these experiments, and then click the Load Images or Load Analytes buttons, respectively.

IsoCode Single-Cell Project Annotation

After your raw data is imported, there will be one row per experiment (chip) in the Project Info table. Each row will contain a set of fields that can be annotated for the corresponding experiment, as seen below.



Below are the available annotation fields for each single-cell sample, from left to right. To start data processing a single-cell experiment, that experiment must have successfully imported images and analytes, and input cell types (including stains). These fields are marked with a double asterisk (**) in the table and in the Project Info table.

Field	Description
Experiment ID	The chip ID (e.g., I-IM123456), along with the date the experiment was created.
Images**	Import status of the IsoLight/IsoSpark raw images from this experiment (a green circle means this was imported successfully).
Analytes**	Import status of the analyte panel information for this experiment/chip (a green circle means this was imported successfully).
Analyst	Name of the analyst processing and analyzing the data.

Notes	Any notes relevant to the experiment.
Cell Type(s)**	Cell sample analyzed in this experiment (e.g., CD8+ cells), along with the stain in parentheses (Red, Blue, or Violet). For example, for CD3+ T cells stained with a Violet Stain use 'CD3 (Violet)'. Use 'Red' or 'Blue' for red and blue stains, respectively. For double-stained samples, type in the second cell type and stain in the second field, e.g., CD8 (Red). Note: double-stained samples should only be run on the IsoLight/IsoSpark if they fall under one of the following two categories: 1. Stain combination has been purchased from Bruker 2. Bruker has verified your stain combination through the IsoPace program. If a stain combination has not been verified by Bruker through one of these avenues, it cannot be guaranteed to work properly and generate accurate single-cell data.
Donor ID	The de-identified donor ID of the analyzed sample, if known.
Donor Groups	One or more distinguishing characteristics of the sample or associated donor in the experiment (e.g., Healthy vs. Diseased, Responder vs. Non-Responder, or more generically Group A vs. Group B vs. Group C). This input can be used later for comparing sample results across patient groups.
Stimulants	The sample stimulant(s), including concentrations and time, if applicable
Viability	The viability of the sample (number ranging from 0 to 1, where 0 = 0% and 1 = 100%). If samples have a wide range of viabilities, these numbers can later be used under Data Analysis for viability normalization of the sample's single-cell data.
Transduction	For re-engineered samples (e.g., CAR-T), the transduction efficiency of the sample (number from 0 to 1, where 0 = 0% and 1 = 100%). If samples have a wide range of transduction efficiencies, these numbers can later be used under Data Analysis for transduction normalization of the sample's single-cell data.

** Indicate required fields to proceed with data processing. Experiments with modified annotations will be marked with an asterisk * in the table. Click the Save button to save all changes to the experiments and project.

Additional Tools and Troubleshooting

New experiments can also be created by clicking the **New** button and specifying the Chip ID of the experiment. The image and analyte information will need to be input for these experiments, using the **Load Images** and **Load Analytes** buttons. If creating new experiments in this fashion, please first move the raw IsoLight/IsoSpark data to the project's Images folder (see Chapter 13 for more information).

Selected experiments can be removed from the project using the **Remove** button.

To add existing experiments back into a project, or move them from a different project, click the Import button, select Removed Experiments, and use the browser to select the .spk files of the experiments you wish to add.

Reset Info will reset any unsaved changes made to the experiments.

A project's annotations can be exported to a csv file by clicking the Export button and selecting Project Info (*.csv).

During an IsoLight or IsoSpark run, if the barcode on one or more chips was not detected, IsoSpeak will query a Chip Barcode database with chip IDs input by the customer, database will return barcodes for corresponding chips, and IsoSpeak will decode barcode into flowsheet information. Please note that internet access is required for this to work. In case of failure of this automated process, please contact support@isoplexis.com to obtain the detailed analyte arrangement for these chips, which will be sent to you in the form of one or more .csv files.

Once received, use the File Explorer on your computer to find the location where your project is saved (e.g., D:\Documents\IsoSpeak Data\My_Project.spkpro)

Save the above .csv files to your project's images folder (e.g., D:\Documents\IsoSpeak Data\My_Project\My_Project_Images\). See Chapter 13 for file organization details.

Chapter 5: Importing CodePlex/Meteor Raw Data and Project Annotation

Once the raw image data is acquired from an instrument run, you will need to import, annotate, and process this data through IsoSpeak to obtain numerical data from the run. The first two steps, outlined in this chapter, involve importing raw data into the IsoSpeak project and then annotating your experiments.

Importing CodePlex Raw Data

To add a new set of experiments that were run on the IsoLight/IsoSpark, go to the **Project Info** module, and click the **Import** button at the top of the screen.

In the subsequent popup, select **Instrument Data**.

Use the browser to select the folder containing the raw IsoLight/IsoSpark data you wish to import and click OK.

The raw data will then get copied to the project's file structure, under the project's **Images** folder (see Chapter 13 IsoSpeak Project File Structure for details).

Note: this step might take a while (for 4 chips, expect ~15 min when importing from a USB 3.0 drive)

Once imported, the new experiments should automatically appear in the table, with their raw images and analyte information already uploaded.

The import status of the raw data is reflected by the circular **Images** and **Analytes** buttons of each experiment (under the **Images** or **Analytes** column in the table).

Experiments under CodePlex projects will have a third **Calibration** button, indicating the import status of the calibration information.

The import status colors are as follows for the buttons:

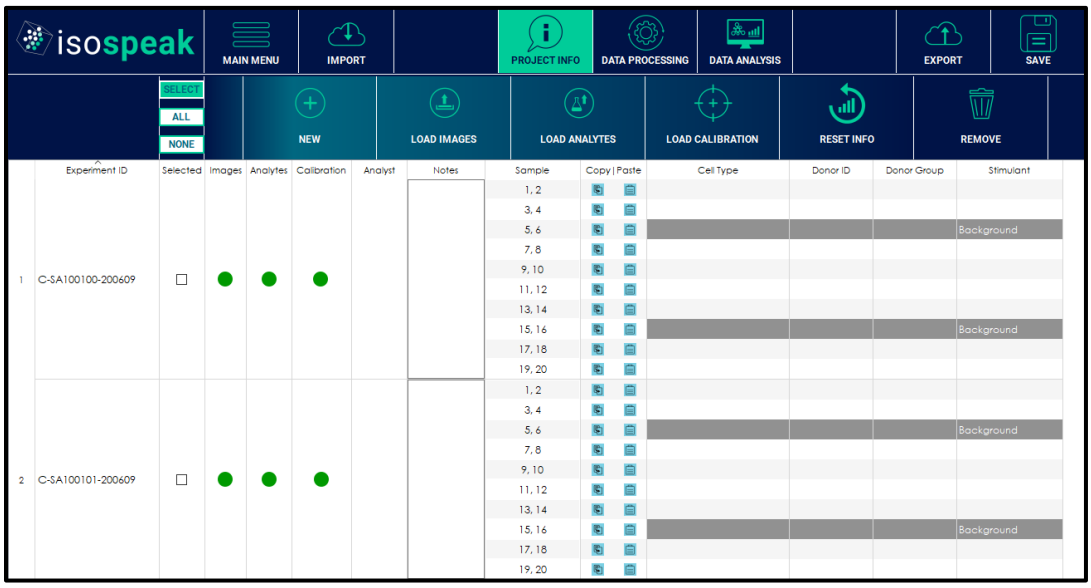
Green: the raw data (images, analytes, or calibration) was loaded successfully.

Blue: the raw data (images, analytes, or calibration) has not yet been input or has not been found.

Red: the previously input raw data (images, analytes, or calibration) cannot be located anymore.

If necessary, raw image data and/or analyte information can be loaded individually after the experiments are created. The data must already reside under the project's **Images** folder for this to succeed (see Chapter 13 IsoSpeak Project File Structure for details).

To load an individual experiment's images or analytes, click the experiment's circular **Images** or **Analytes** button. To batch-load images or analytes for multiple experiments, select these experiments, and then click the **Load Images** or **Load Analytes** buttons, respectively.



CodePlex Project Annotation

After your raw data is imported for a CodePlex project, there will be one row per experiment (chip) in the **Project Info** table. The left portion of each row will contain a set of fields that can be annotated for the corresponding experiment. Additionally, since CodePlex chips can have multiple samples, there will be several fields on the right that correspond to specific samples of a chip. See the figure above.

Below are the available **chip-level** annotation fields for each CodePlex experiment. **To process the data of a CodePlex experiment, it must have successfully imported images, analytes, and calibration data.**

Field	Description
Experiment ID	The chip ID (e.g., I-IM123456), along with the date the experiment was created.
Images	Import status of the IsoLight/IsoSpark raw images from this experiment (a green circle means this was imported successfully).
Analytes	Import status of the analyte panel information for this experiment/chip (a green circle means this was imported successfully).
Calibration	Import status of the calibration chip information for this experiment/chip (a green circle means this was imported successfully)
Analyst	Name of the analyst processing.
Notes	Any notes relevant to the experiment.

There are additional annotation fields available for each CodePlex **sample**. Below are the available **sample-level** annotation fields.

While no sample-level fields are required to proceed with data processing, samples with no annotations will not be included in the final data.

Field	Description
Sample	Noneditable field with the sample label, corresponding to the indices of the two wells that the sample was loaded into on the CodePlex chip.
Copy | Paste	Buttons that allow you to copy all editable annotations from one sample and paste them into the fields of another sample. Clicking the **Copy** (left) button next to a sample copies the annotations of this sample and highlights the button orange. Clicking the **Paste** (right) button next to a different sample will then paste the copied annotations into this sample's fields.
Cell Types	Cell type of this sample. Unlike in single-cell experiments, you do not need to type in any stain information (not applicable) – only the Cell Type (e.g., CD8+ cells).
Donor ID	The de-identified donor ID of the analyzed sample, if known.
Donor Groups	One or more distinguishing characteristics of the sample or associated donor in the experiment (e.g., Healthy vs. Diseased, Responder vs. Non-Responder, or more generically Group A vs. Group B vs. Group C). This input can be used later for comparing sample results across patient groups.
Stimulants	The sample stimulant(s), including concentrations and time, if applicable

Experiments with modified chip- or sample-level annotations will be marked with an asterisk *. Click the Save button to save all changes to the experiments and project.

Note that CodePlex annotations corresponding to **Background** samples (samples 5, 6 and samples 15, 16) are grayed out and not editable.

Importing Meteor Raw Data

To add a new set of experiments that were run on the IsoSpark, go to the **Project Info** module, and click the **Import** button at the top of the screen.

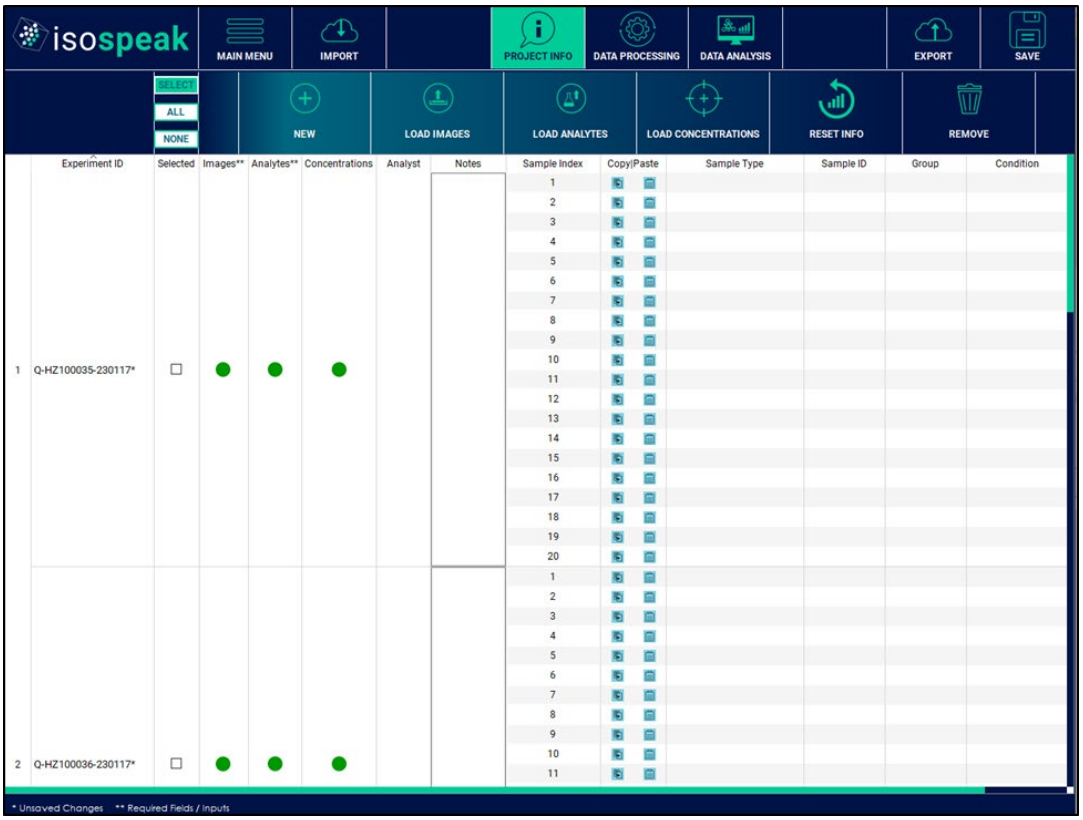
In the subsequent popup, select **Instrument Data**.

Use the browser to select the folder containing the raw IsoLight/IsoSpark data you wish to import and click OK.

The raw data will then get copied to the project's file structure, under the project's Images folder (see Chapter 13 IsoSpeak Project File Structure for details).

Note: this step might take a while (for 4 chips expect ~15 min when importing from a USB 3.0 drive)

Once imported, the new experiments should automatically appear in the table, with their raw images and analyte information already uploaded.



The import status of the raw data is reflected by the circular Images and Analytes buttons of each experiment (under the Images or Analytes column in the table). Note that Concentrations are loaded separately (see next section).

The import status colors are as follows for the buttons:

Green: the raw data (images, analytes, or concentrations) was loaded successfully.

Blue: the raw data (images, analytes, or concentrations) has not yet been input or has not been found.

Red: the previously input raw data (images, analytes, or concentrations) cannot be located anymore.

If necessary, raw image data and/or analyte information can be loaded individually after the experiments are created. The data must already reside under the project's Images folder for this to succeed (see Chapter 13 IsoSpeak Project File Structure for details).

To load an individual experiment's images, analytes, or concentrations, click the experiment's circular Images, Analytes, or Concentrations button. To batch-load images or analytes for multiple experiments,

select these experiments, and then click the Load Images, Load Analytes, or Load Concentrations buttons, respectively.

Inputting Quantitation Information for Meteor Data

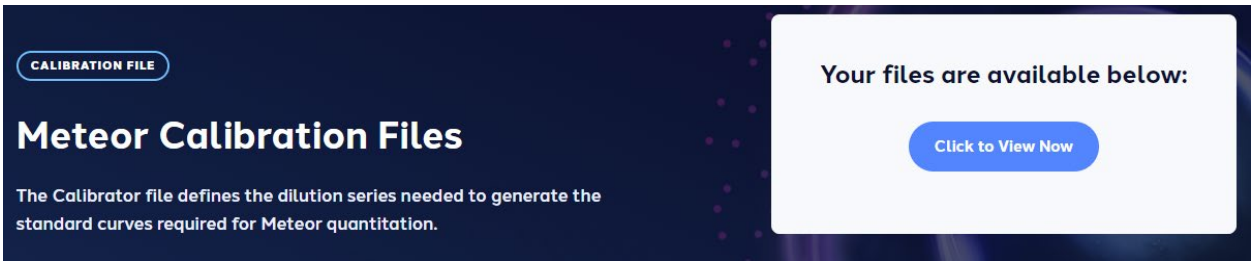
To obtain quantified data from your samples, in addition to importing the raw data there are two pieces of information required by IsoSpeak:

1. Loading in the appropriate calibrator concentrations file, and
2. Annotating which samples are calibrator and background samples

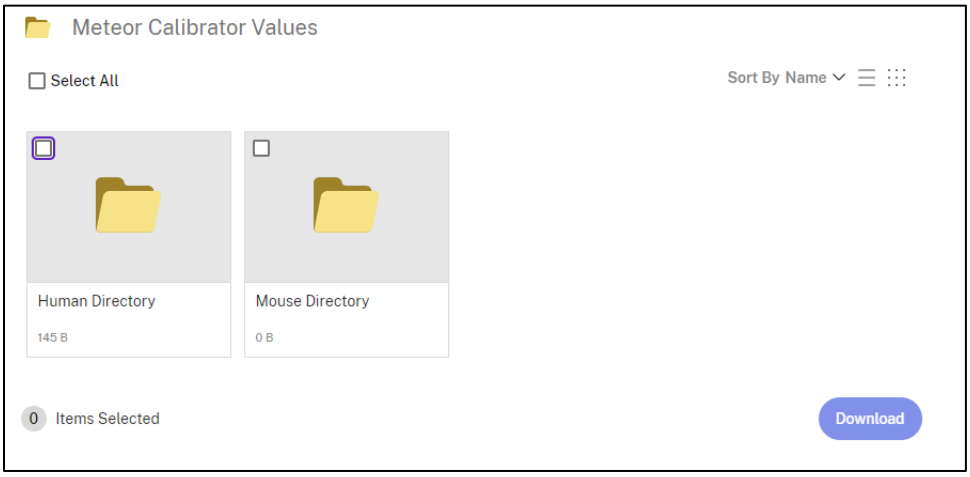
These two inputs are required to obtain quantified (pg/ml concentration) data.

For details on how quantitation is performed, please see the Meteor Calibration Curves section.

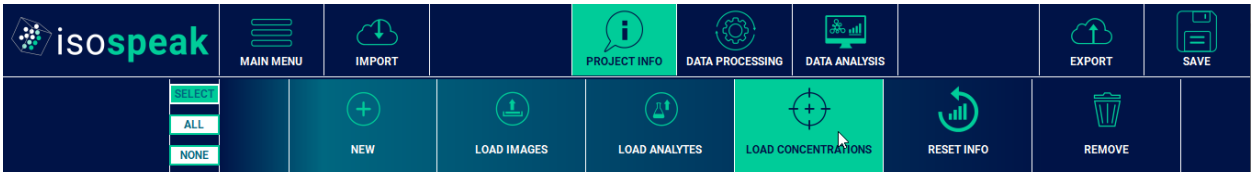
To load the calibrator concentrations file corresponding to your chips into IsoSpeak, you will first need to obtain this file by going to https://phenomex.com/resource/meteor-calibration-files/.



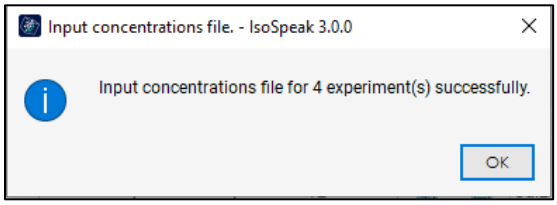
Navigate to the relevant folder (Human or Mouse) and download either the entire folder or the individual file whose chip range includes the chip IDs in your instrument run (e.g., if your chips are Q-HZ100035 and Q-HZ100036, download the file with the chip range Q-HZ100000-HZ100200.)



Once you have downloaded the file to your computer, go to IsoSpeak, select the chips/experiments (individually or Select → All), then click the Load Concentrations, and select the file you just downloaded using the file browser, and click Ok.



If the correct file is selected, you should then see a popup like the one below, and the Concentrations buttons of the selected chips should turn green.



Next, you will need to annotate which of your samples are Calibrator and Background Calibrator samples using the Sample Type field, so IsoSpeak knows which samples' data to use for generating a calibration curve for quantitation. The available options in this field are:

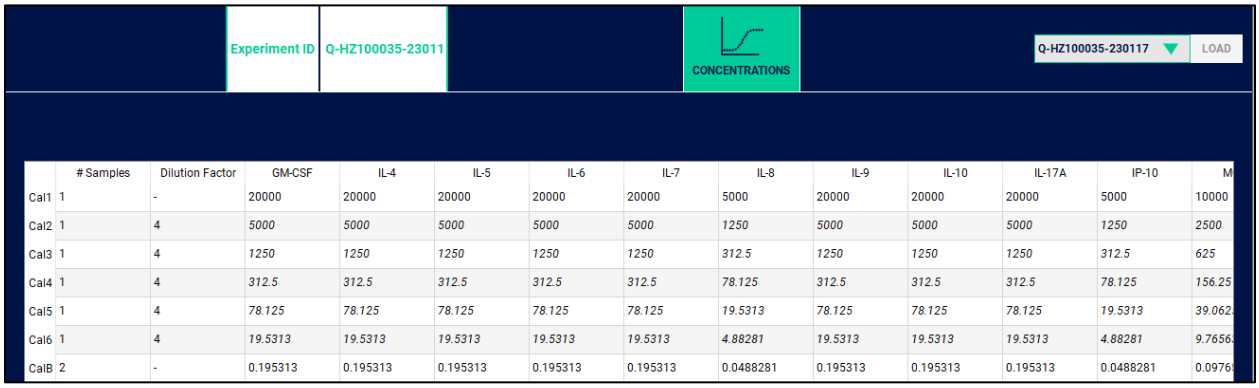
- Cal1, Cal2, Cal3, … – specifies a sample is a calibrator sample, with the number indicating the calibrator level (1 = highest concentration, 2 = second highest concentration, etc.).
- CalB – background calibrator samples, optionally used in generating the calibration curve (data can be included or not under Data Analysis).
- [Blank] – leave this field blank if the corresponding sample is not a calibrator or background sample, to ensure its data is excluded from calibration curve fitting.

Data from Calibrator samples will be used for calibration curve generation, and Background samples are optionally used (option under Data Analysis). IsoSpeak requires a minimum of 6 calibrator levels (one sample annotated Cal1 through Cal6). Two background calibrator samples (CalB) are recommended, and these are optionally used for curve fitting (toggle under Data Analysis).

Experiment ID	Selected	Images**	Analytes**	Concentration	Analyst	Notes	Sample Index	Copy|Paste	Sample Type
							1		
							2		Cal4
							3		
							4		
							5		Cal5
							6		
							7		
							8		Cal6
							9		Cal3
							10		
Q-HZ100035-230117	☐	●	●	●			11		
							12		Cal2
							13		
							14		
							15		CalB
							16		
							17		
							18		
							19		Cal1
							20		CalB

Once you have input both pieces of information (calibration concentrations file and sample type annotations, you may verify that IsoSpeak has all the required information needed to generate calibrations

curves to quantify your sample data. Click the green Concentrations button of a chip/experiment and click View in the resulting popup to see the input quantitation data.



	# Samples	Dilution Factor	GM-CSF	IL-4	IL-5	IL-6	IL-7	IL-8	IL-9	IL-10	IL-17A	IP-10	M
Cal1	1	-	20000	20000	20000	20000	20000	5000	20000	20000	20000	5000	10000
Cal2	1	4	5000	5000	5000	5000	5000	1250	5000	5000	5000	1250	2500
Cal3	1	4	1250	1250	1250	1250	1250	312.5	1250	1250	1250	312.5	625
Cal4	1	4	312.5	312.5	312.5	312.5	312.5	78.125	312.5	312.5	312.5	78.125	156.25
Cal5	1	4	78.125	78.125	78.125	78.125	78.125	19.5313	78.125	78.125	78.125	19.5313	39.062
Cal6	1	4	19.5313	19.5313	19.5313	19.5313	19.5313	4.88281	19.5313	19.5313	19.5313	4.88281	9.7656
CalB	2	-	0.195313	0.195313	0.195313	0.195313	0.195313	0.0488281	0.195313	0.195313	0.195313	0.0488281	0.0976

For the given chip, you will see a table of all calibrator levels and backgrounds, the number of samples of each, and the concentrations at each calibrator level and for each analyte.

You may see the following warnings here if the quantitation information was not properly input:

- Not all calibrator concentrations have been defined for all proteins. Please make sure the correct concentrations file was loaded in under Project Info.
- Fewer than 6 calibrators have been defined, the minimum required for calibration curve creation. Please make sure the calibrator samples have been identified using the Sample Type field under Project Info

Once inside the calibration information for an experiment, you can use the Experiment/Chip drop down in the upper right corner to switch in turn through all the experiments/chips.

Meteor Project Annotation

After raw data is imported for a Meteor project, there will be one row per experiment (chip) in the Project Info table. The left portion of each row will contain a set of fields that can be annotated for the corresponding experiment. Additionally, since Meteor chips contain multiple samples, there will be several fields on the right that correspond to specific samples of a chip. See the figure below.

Below are the available chip-level annotation fields for each Meteor experiment. To process the data of a Meteor experiment, it must have successfully imported images, analytes, and calibration data.

Field	Description
Experiment ID	The chip ID (e.g., I-IM123456), along with the date the experiment was created.
Images	Import status of the IsoLight/IsoSpark raw images from this experiment (a green circle means this was imported successfully).
Analytes	Import status of the analyte panel information for this experiment/chip (a green circle means this was imported successfully).

Concentrations	Import status of the calibrator concentration information for this experiment/chip (a green circle means this was imported successfully)
Analyst	Name of the analyst processing.
Notes	Any notes relevant to the experiment.

There are additional annotation fields available for each Meteor **sample**. Below are the available **sample-level** annotation fields.

No sample-level fields are required to proceed with data processing; however, **the Sample Type information (along with inputting the calibrator concentration file) is required to generate calibration curves and obtain quantified data** (see previous section).

Field	Description
Sample Index	Noneditable field with the sample label, corresponding to the index of the well that the sample was loaded into on the Meteor chip.
Copy | Paste	Buttons that allow you to copy all editable annotations from one sample and paste them into the fields of another sample. Clicking the **Copy** (left) button next to a sample copies the annotations of this sample and highlights the button orange. Clicking the **Paste** (right) button next to a different sample will then paste the copied annotations into this sample's fields.
Sample Type	The type of sample (Calibrator, Background, or blank to indicate neither). This annotation is relevant for quantifying your data – data from Calibrator and Background samples will be used to generate a calibration curve for quantification. **This input is required to obtain quantified (pg/ml concentration) data.** The available options are: • **Cal1, Cal2, Cal3, …** – specifies a sample is a calibrator sample, with the number indicating the calibrator level (1 = highest concentration, 2 = second highest concentration, etc.). • **CalB** – background/empty wells, optionally used in generating the calibration curve (data can be included or not under Data Analysis). • **[Blank]** – leave this field blank if the corresponding sample is not a calibrator or background sample, to ensure its data is excluded from calibration curve fitting.
Sample ID	The sample ID of the analyzed sample, if known.
Group	One or more distinguishing characteristics of the sample or associated donor in the experiment (e.g., Healthy vs. Diseased, Responder vs. Non-Responder, or more generically Group A vs. Group B vs. Group C). This input can be used later for comparing sample results across patient groups.
Condition	The sample condition(s), if known

Experiments with modified chip- or sample-level annotations will be marked with an asterisk *. Click the Save button to save all changes to the experiments and project.

Additional Tools and Troubleshooting

New experiments can also be created by clicking the New button and specifying the Chip ID of the experiment. The image and analyte information will need to be input for these experiments, using the Load Images and Load Analytes buttons. To quantify data, the calibration information will also need to be input for CodePlex experiments, using the Load Calibration button, whereas the calibrator concentration information will need to be input for Meteor experiments, using the Load Concentrations button. If creating new experiments in this fashion, please first move the raw IsoLight/IsoSpark data to the project's Images folder (see Chapter 13 for more information).

Selected experiments can be removed from the project using the Remove button.

To add existing experiments back into a project, or move them from a different project, click the Import button, select Removed Experiments, and use the browser to select the .spk files of the experiments you wish to add.

Reset Info will reset any unsaved changes made to the experiments.

A project's annotations can be exported to a csv file by clicking the Export button and selecting Project Info (*.csv).

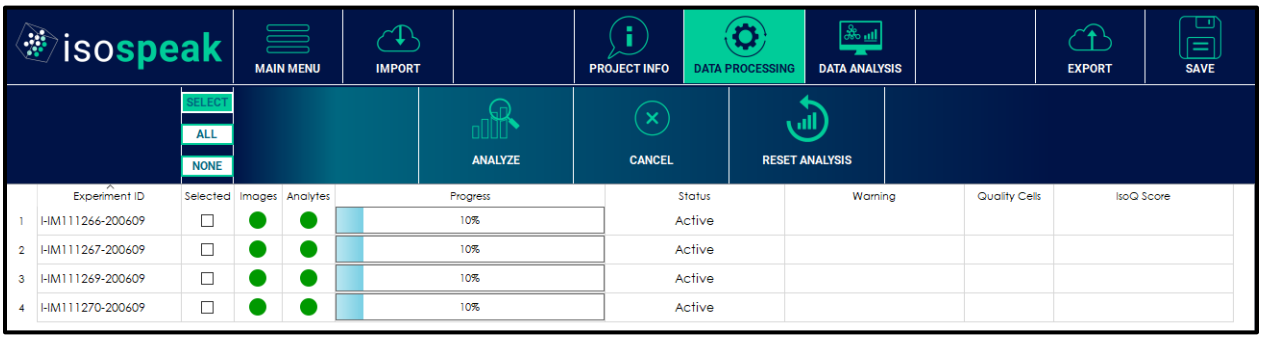
During an IsoLight or IsoSpark run, if the barcode on one or more chips was not detected, IsoSpeak will query a Chip Barcode database with chip IDs input by the customer, database will return barcodes for corresponding chips, and IsoSpeak will decode barcode into flowsheet information. Please note that internet access is required for this to work. In case of failure of this automated process, please contact support@isoplexis.com to obtain the detailed analyte arrangement for these chips, which will be sent to you in the form of one or more .csv files.

Once you receive csv files, use the File Explorer on your computer to find the location where your project is saved (e.g., D:\Documents\IsoSpeak Data\My_Project.spkpro) and save the csv files to your project's images folder (e.g., D:\Documents\IsoSpeak Data\My_Project\My_Project_Images\). Bring up IsoSpeak and the csv files should become part of the experiment data. See Chapter 13 for file organization details.

Once the IsoLight/IsoSpark raw data has been imported and annotated, the next step is automated data processing of this data, to go from raw image data to final numerical results. Along the way, the data undergoes rigorous quality control to flag any potential issues that could affect the accuracy of the data.

IsoCode Data Processing

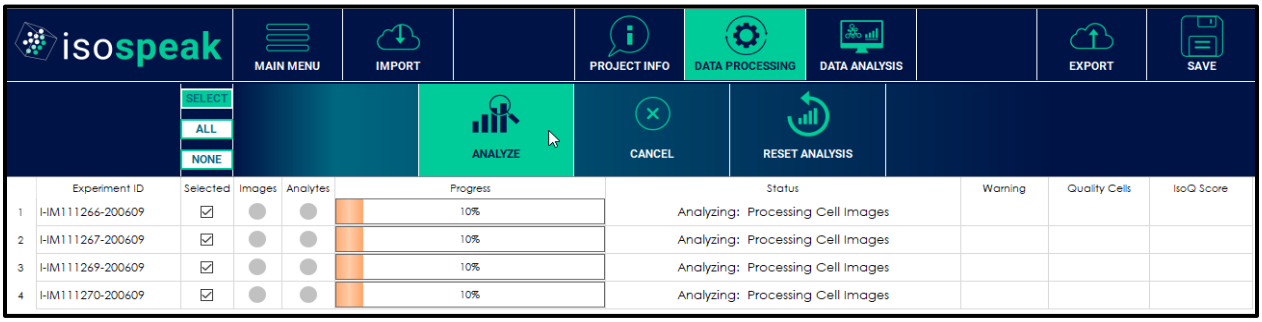
Click the Data Processing button to start the automated analysis.



Before starting the analysis, ensure that the Images and Analytes buttons are both green for each experiment (raw data has been imported successfully).

For Single-Cell projects, also make sure that the cell type field of each experiment has been populated under the Project Info module.

Select the experiments you wish to process, and click the Analyze button to begin data processing:

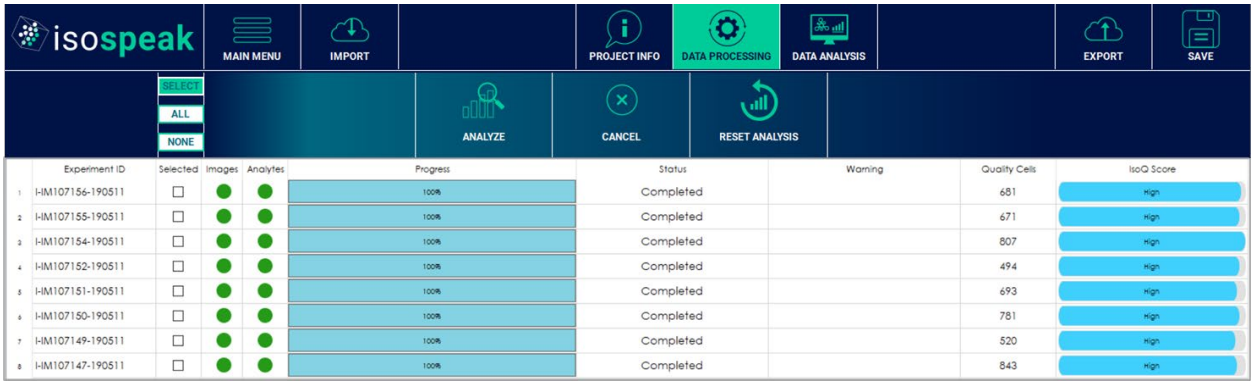


A progress bar indicates the current stage of the data processing. Analysis details will appear to the right of the progress bar.

Once data processing is completed, you can use the tools available in Data Analysis to analyze the data in this project (see Chapter 8).

IsoCode Single-Cell Data Quality Control

The IsoQ Score, found in the Data Processing module next to each processed sample, is generated during processing to indicate the quality of the conversion to results. Samples are analyzed for cell quality during this phase and given an IsoQ score. All samples should have a blue IsoQ score indicating that the sample quality passes, and the data can proceed through the automated IsoSpeak analysis. A dark purple score with the text "Low" indicates that there is a possible detected issue. The number of Quality Cells in the field next to the IsoQ score indicates the number of quality single cells that were included in the final data for the corresponding experiment.

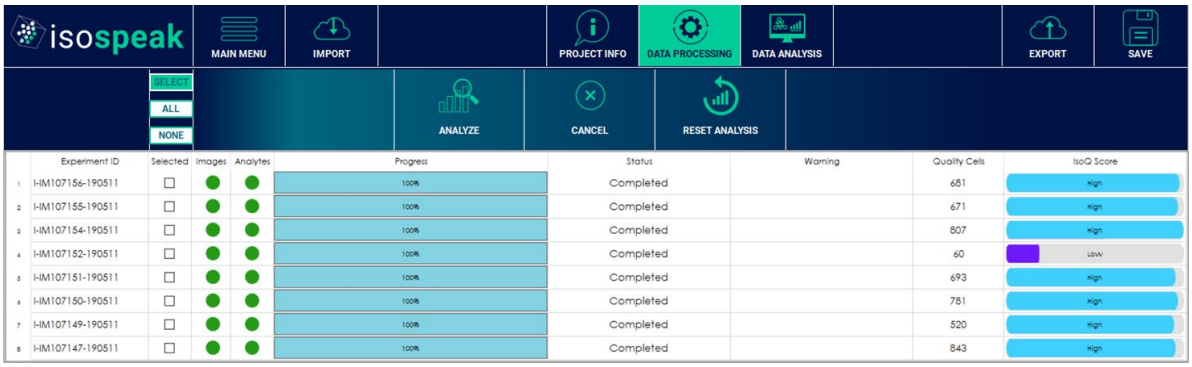


	Experiment ID	Selected	Images	Analytes	Progress	Status	Warning	Quality Cells	IsoQ Score
1	I-IM107156-190511	☐			100%	Completed		681	High
2	I-IM107155-190511	☐			100%	Completed		671	High
3	I-IM107154-190511	☐			100%	Completed		807	High
4	I-IM107152-190511	☐			100%	Completed		494	High
5	I-IM107151-190511	☐			100%	Completed		693	High
6	I-IM107150-190511	☐			100%	Completed		781	High
7	I-IM107149-190511	☐			100%	Completed		520	High
8	I-IM107147-190511	☐			100%	Completed		843	High

Troubleshooting. The IsoQ score should be used to judge the success of sample analysis meeting known given condition metrics.. The quality metric considers a combination of sample-related elements required to meet specifications, per Bruker protocols, that indicate quality single-cell proteomic data. The QC metrics gathered and combined are presented as a final score called the IsoQ score and is color-code to indicate meeting or exceeding a total threshold. Success is indicated by blue high values. Issues are indicated by low purple values.

The components of the IsoQ score correlate to effective sample staining and cell counts, expected background levels and consistency within the signal data, and obtaining accurate and sample-specific data for each measured protein in an automated fashion. If a purple IsoQ score is generated, it may indicate one of the following issues:

- Low cell counts that affect the statistical significance of the data, possibly a result of cell staining or cell loading issues.
- Substantial artifacts identified around the imaged signal information, potentially affecting the quality of the resulting data.
- Significant variation in the background noise levels across the signal image, potentially affecting the sensitivity of the resulting data.
- Ascertaining quality single-cell specific protein data



Bruker may ask you to upload your processed experiments to the Citrix ShareFile folder created for you by Bruker to further quality control your data and verify its integrity. The quality-controlled data will then be transferred back to ShareFile for you to download and analyze. For details, see Chapter 8 Uploading Data to Citrix ShareFile for Quality Analysis. The following table provides troubleshooting recommendations for areas that can contribute to a Purple IsoQ score, such as improving low quality cell counts, cell staining and techniques to improve background noise and reduction of debris.

Contributor to low Q Score	Possible Reason	Recommended Solutions
Low quality cell count on chip ***Cell counting & concentration related***	• Recommended cell concentrations not used • Issue with Cell Counting procedure • Trypan Blue may have debris • Poor cell removal from cell culture plate	• Use recommended cell concentrations during overnight incubation • Use appropriate dilutions recommended • Quick spin Trypan Blue to pellet potential debris, remove aliquot from top of Trypan Blue. • Start with fresh aliquot of Trypan Blue • Thoroughly mix cells in well with pipette prior to transferring to tube
Low quality cell count on chip ***Stain process related***	• Use of media other than the recommended media in protocol which could interact with cell stain • Use of stains not recommended in protocol • Recommended stain concentration, incubation time and/or incubation temperature not used • Cell stain was stored prior to use	• Use complete RPMI media • Use IsoPlexis provided validated stain • Follow recommended staining steps • Use only freshly prepared membrane stain

Low quality cell count on chip
Technique detail related
	• Bubbles loaded onto chip, especially at Chip Loading
	• Detection of potential artifacts such as debris, cell clumping, inefficient enrichment possibly due to:
	  • Pipetting wrong concentration
	  • Not fully inserting column into MACS separator
	  • Reagents not stored at recommended temperatures
	• Recommended number of cells not loaded on chip
		• Avoid introduction of bubbles on chip by mixing and pipetting carefully
		• Ensure use of a sterile space to reduce introduction of potential contaminants.
		• Use dedicated pipettes, tips, and tubes for sterile work. Pipette up and down gently and throughout protocol to reduce clumps.
		• Load recommended number of cells (30,000 cells per chip)

Limited frequency of stimulated cells, i.e., those with cytokine signal ***Viability related***	• Leaving thawed cells in DMSO for an extended period • Low viable cells due to low viability input sample and lack of utilization of Ficoll-Paque • Decreased viability due to cell shock	• After thaw, quickly transfer cells from DMSO to complete RPMI to ensure viability of cells. • Verify viability of cells is above 80% • Use reagents at recommended temperatures (i.e. always use warmed media [37°C])
Limited frequency of stimulated cells, i.e., those with cytokine signal ***Stimulation step related***	• Recommended stimulation concentration was not used • Recommended stimulation duration was not used	• Use IsoPlexis recommended stimulation concentrations • Follow IsoPlexis recommended stimulation times • Use IsoPlexis recommended reagents

Additional Tools and Troubleshooting

To stop the analysis of one or more experiments, select them and click the **Cancel** button. The experiments in progress will continue running until the current analysis step (e.g., cell detection) is completed.

Use the **Reset Analysis** button if you wish to rerun the analysis of an experiments (for example, if you re-imported the raw data inputs and wish to restart the analysis from the beginning).

In case of the following error messages, please go back to the **Project info** module and re-input or re-import the correct experiment information:

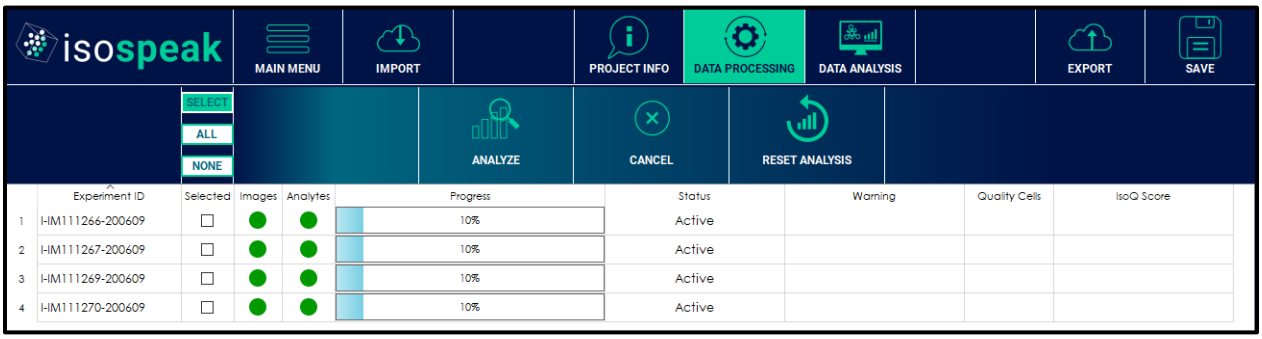
1. Missing cell or signal images: please import raw IsoLight/IsoSpark images.
2. Missing analytes: please import analyte information.
3. Missing cell types/stains (only for single-cell projects): please input the analyzed cell samples and corresponding stains.

For other errors of the type **Processing failed with Error Code X**, please contact support@isoplexis.com for assistance.

Once the IsoLight/IsoSpark raw data has been imported and annotated, the next step is automated data processing of this data, to go from raw image data to final numerical results. Along the way, the data undergoes rigorous quality control to flag any potential issues that could affect the accuracy of the data.

CodePlex and Meteor Data Processing

Click the Data Processing button to start the automated analysis.

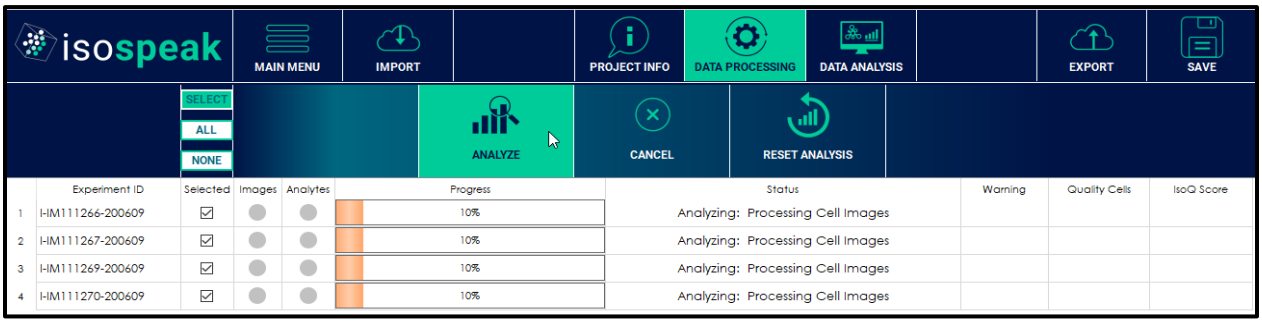


Before starting the analysis, ensure that the Images and Analytes buttons are both green for each experiment (raw data has been imported successfully).

For CodePlex projects, ensure that the Calibration buttons are also green for each experiment. Additionally, make sure you have added at least one annotation to all your samples under the Project Info module, otherwise they will be excluded in the final analysis.

For Meteor projects, ensure that the Concentrations buttons are also green for each experiment.

Select the experiments you wish to process, and click the Analyze button to begin data processing:

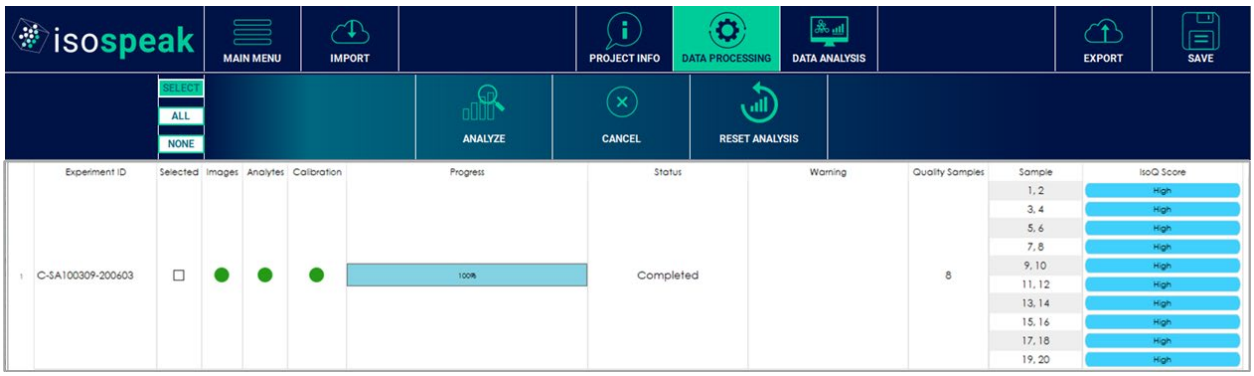


A progress bar indicates the current stage of the data processing. Analysis details will appear to the right of the progress bar.

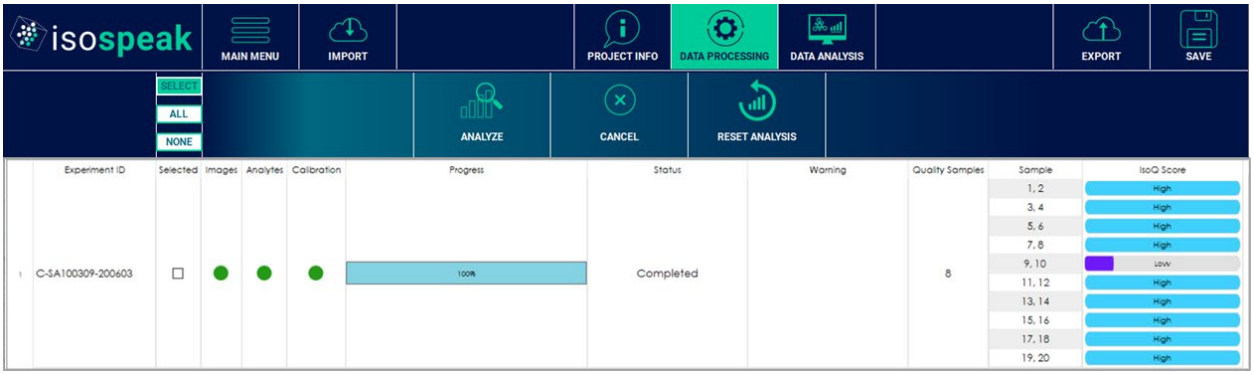
Once data processing is completed, you can use the tools available in Data Analysis to analyze the data in this project (see Chapter 10).

CodePlex Data Quality Control

The IsoQ score for CodePlex samples is an automated quality metric for each of the samples processed through IsoSpeak. An IsoQ score is displayed for each sample in each CodePlex chip (including background samples), i.e. there will be a total of ten IsoQ scores generated by the software for every chip. Samples are analyzed for intra- and inter-well consistency and issues with analyte signal saturation during the data processing phase and given a cumulative IsoQ score. Blue IsoQ scores are normal and indicate high sample and data quality. The number of samples included in the final data is indicated in the Quality Samples field. This is typically eight samples, if this is how many non-background samples were originally loaded onto the chip.



Troubleshooting. The IsoQ score is an indicator of the reliability and accuracy of the data from a particular sample. The components of the CodePlex IsoQ score correlate to a sample's intra- and inter-well consistency and any issues with analyte signal saturation. If a reported IsoQ score is purple, there may be an issue with the consistency or signal saturation.



As a general practice, Bruker will assist with troubleshooting all runs with reported purple IsoQ scores.

However, if for any reason you are uncertain or concerned about the quality of your data, Bruker can assist in confirming its validity and accuracy.

Users should export a log of the IsoQ scores by going to Export -> IsoQ Score Log and send the log to Bruker customer support for further assistance.

Bruker may also ask you to upload your processed experiments to the Citrix ShareFile folder created for you by Bruker to further quality control your data and verify its integrity. The quality-controlled data will then be transferred back to ShareFile for you to download and analyze. For details, see Chapter 13 Uploading Data to Citrix ShareFile for Quality Analysis.

Meteor Data Quality Control

Currently, no IsoQ score is available for Meteor samples. However, there are several steps you can take to verify the quality of your data:

1. Review statistical information / variability of the intra-sample replicates (see Meteor Raw Data Table section for details).
2. Review statistical information / variability of any inter-sample replicates, if replicate samples were included in the run (see Meteor Signal Intensity Table section for details).
3. If quantifying data, check whether calibration curves were successfully generated, and there are no curve generation errors present (see Meteor Standard Curves section for details).
4. If curve generation failed (curve parameters absent in Standard Curves table and "No Curve" present in data tables):
    a. Confirm that you have correctly input the calibrator sample information (see Inputting Quantitation Information for Meteor Data section for details).
    b. Try to generate curves without the inclusion of Background Calibrator samples (deselect Format Options → Include Background Calibrators in Curve Fitting) – in rare cases these data points can negatively impact the successful generation of calibration curves.
5. If quantifying data, review goodness of fit metrics for generated calibration curves. Additionally, use Bottom and Top parameters of the curve (in RFU) and LLOD and UL metrics (in pg/ml) to confirm that sample intensities can effectively be quantified, i.e., intensities are within those ranges (see Meteor Standard Curves section for details).

Bruker may also ask you to upload your processed experiments to the Citrix ShareFile folder created for you by Bruker to further quality control your data and verify its integrity. The quality-controlled data will then be transferred back to ShareFile for you to download and analyze. For details, see Chapter 13 Uploading Data to Citrix ShareFile for Quality Analysis.

Additional Tools and Troubleshooting

To stop the analysis of one or more experiments, select them and click the Cancel button. The experiments in progress will continue running until the current analysis step (e.g., cell detection) is completed.

Use the Reset Analysis button if you wish to rerun the analysis of experiments (for example, if you re-imported the raw data inputs and wish to restart the analysis from the beginning).

In case of the following error messages, please go back to the Project info module and re-input or re-import the correct experiment information:

1. Missing signal images: please import raw IsoLight/IsoSpark images.
2. Missing analytes: please import analyte information.

For other errors of the type Processing failed with Error Code X, please contact support@isoplexis.com for assistance.

Chapter 8: IsoCode Single-Cell Data Analysis Overview

Once you have processed and obtained data for one or more experiments, you may use the data analysis tools provided in IsoSpeak to begin analyzing the experimental results. You have the option of looking at as many or as few experiments as you want, and experiments can come from multiple instrument runs or other projects of the same chip type.

Data Analysis Interface

From the Main Menu or the Main Toolbar, click the Data Analysis button. You will see this interface.

The Main Toolbar at the top lets you go to different program stages, as well as go back to the Main Menu.

You can also use the Export button in the Main Toolbar at any time to export the currently displayed graphs or table. Graphs are output as files of extension png, while tables are output as tab-delimited text files of extension txt or as comma-delimited text files of extension csv.

The Select Data Type toolbar at the top and the Select Table/Chart sub-menu let you select your desired visualization (or table), which will show up beneath these toolbars. A horizontal scrollbar below the Select Table/Chart sub-menu allows you to scroll left or right to easily access all available options in both Select Data Type and Select Table/Chart toolbars.

On the left, you can toggle between Filter Selection options and Format Options.

The Filter Selection options allow you to select specific subsets of the data from the loaded experiment(s). You can filter the data based on stimulation conditions, experiments, donor IDs, donor groups, cell markers, and proteins. Adjusting the filters will automatically recreate the displayed graph or table, only showing the samples and sample subsets meeting the criteria of the selected filters. You can also collapse the options that you are not using by clicking on the green triangle to the left of the option headers. Scroll down the list of filter selections using the vertical scrollbar if the selection list is too long to fit on your screen.

An additional capability of the Filter Selection options is the ability to find average data across stimulations, donors, donor groups, etc. For example, if you profiled 5 donors with the same stimulation conditions and you wish to know the average profile of these donors, simply click the Average button under Donors, and the displayed table or graph will show the average profile. You can average across any number of criteria at once (for example, average across Donors and Markers simultaneously). If any of the graphs are averages of multiple samples (see Filter Selection above), error bars will typically be shown by default to illustrate the variability across samples.

Under Filter Selection, General you will also find chip-family specific options such as Normalize Data for Single-Cell Signaling datasets (see Chapter 9 for details).

Format Options provides various options for fine-tuning the properties, layout, and look of the displayed graphs and tables. For example, you can select how samples are sorted, what gets displayed as the sample title, and so on.

For any of the displayed graphs, you can use the scroll wheel on your mouse to zoom in an out. Alternatively, you can specify the zoom amount at the top right of the screen.

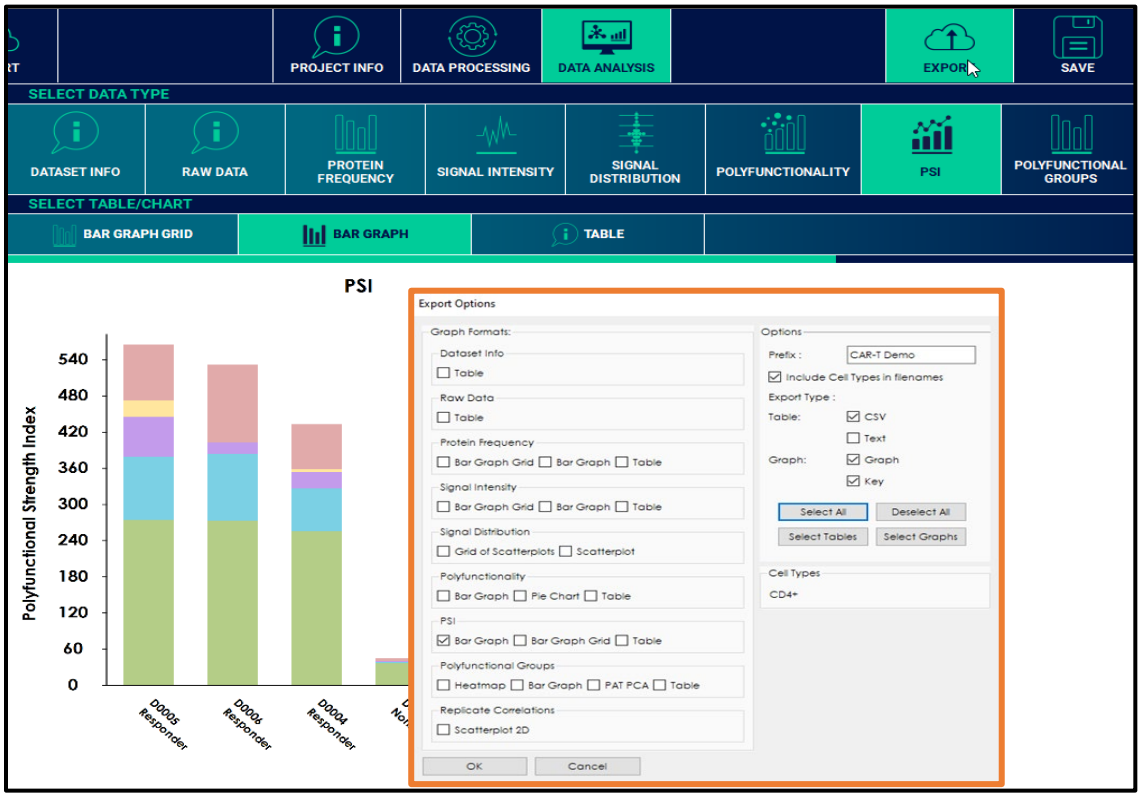
Some of the single-cell visualizations will color-code individual proteins according to the functional group they belong to, where each functional group represents the general role of the protein (e.g., Effector or Stimulatory). The functional groups and associated colors are pre-defined and specific to each single-cell chip (although some chips with similar protein panels may also have the same set of functional groups).

Some visualizations are exclusive to specific project types (chip families). See Chapter 9 for the visualizations available for this chip family.

Exporting Data

To export any data or visualization from the Data Analysis module, click on the Export button. You will see an Export Options dialog pop up, with the selected option matching the currently visible graph/table.

To export screenshots of 3D graphs (e.g., t-SNE), you will need to go to this graph type, rotate the graph to the desired view, and then click the button at the top-right of the 3D graph viewer.



The Export Options dialog has several options. On the left, under Graph Formats, users can choose from different available table or graph types and can check any of the items to be exported at the same time. IsoSpeak will export each selected table or graph to a file based on your selection in the Options area.

Options allows you to define the prefix and export type of each graph/table, and whether cell types are included in the exported filenames. The default filename prefix is the project name. Cell types are included by default in the filenames. Under export type, users can choose between "CSV" or "Text" formats when exporting tables, and "Graph" or "Key" when exporting graphs.

Cell Types displays the cell samples currently being viewed.

Chapter 9: IsoCode Single-Cell Data Analysis Table and Graph Details

This section describes the data types that can be explored in IsoSpeak for a given dataset, along with the various ways in which each data type can be viewed. Please note that some data types or display types are only available for certain projects.

Dataset Info

Table

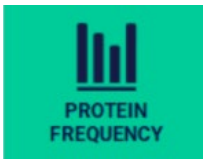
Displays general metadata about the experiments, such as the donor IDs, the cell samples, the donor groups, and so on. Much of this is what was previously input under Project Info.

Raw Data

Table

For Single-Cell projects, this table shows the raw single-cell data obtained for the selected experiment. Each row corresponds to a single cell from one of the experiments, and the columns indicate the signal intensity of each secreted cytokine by the cell.

Protein Frequency



Bar Graph Grid

This set of bar graphs shows the percentage of single cells (y-axis) in each sample that secrete each of the selected analytes (x-axis). The color-coding of the analytes corresponds to the functional group they are assigned to (e.g., for Single-Cell Adaptive Immune: Effector, Stimulatory, Inflammatory, Regulatory, or Chemoattractive). Bars that are gray correspond to analytes that were not significantly secreted by the sample (by default, less than 2% of single cells).

Bar Graph

This bar graph shows the percentage of single cells (y-axis) in each sample that secrete each of the selected analytes (x-axis). The bars are color-coded based on the sample.

Table

This table shows the percentage of single cells in each sample (rows) that secrete each of the selected analytes (columns).

Signal Intensity

Bar Graph Grid

This set of bar graphs shows the average observed intensities (y-axis) in each sample for each of the selected analytes (x-axis). If samples have been averaged into groups, the bar graph will show the average signal intensity (y-axis) of each analyte (x-axis) of the sample group. Under Format Options, you can choose whether the displayed data is Raw, Thresholded (background subtracted), Log Transformed, or Log Thresholded (thresholded and log transformed).

Bar Graph

This bar graph is equivalent to the above set of bar graphs, but presents the data in a single graph, with bars color-coded based on the sample.

Heat Map (Single-Cell Signaling)

The Signal Intensity Heat Map is only present with Single-Cell Signaling Projects. This heat map displays differences in observed protein signal intensities across samples. Each square of the heat map represents the average signal intensity of a given protein in each sample or sample group. Higher signal intensities are yellow, while lower signal intensities are purple.

The sample data is organized in rows and columns, with the first Sort By... option under Format Options corresponding to the factor along the columns, and the second corresponding to the factor along the rows. The sample labels on the left are determined by the Label By... options.

By default, the presented data is log transformed; this may be turned off under Format Options. If the Normalize Data button is selected under Filter Selection, the data is also first normalized to the average signal intensity of the housekeeping control protein (Alpha Tubulin).

By default, the colors are not scaled (No Scaling option). The colors can further be scaled per column of data in several ways:

- User-Specified scaling will scale to the values input in the Min. Value and Max. Value options (purple will correspond to the min. value and any value lower than that, while yellow will correspond to the max. value and any value higher than that)
- Standard scaling will divide the data values by the standard deviation of each column of data (Note: will not be applied to quantified data)

- Normalization (0-1) will subtract the minimum data value of the column, and then divide by the range of the data (maximum value – minimum value) in the column (Note: will not be applied to quantified data)

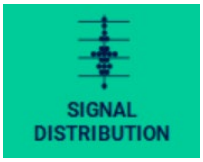
Activating Format Options → Separate Sample Groups will separate and label each set of rows of the heatmap, according to the first Sort By… property (e.g., if you sort by Experiment ID, then each experiment's samples will be separate in the heat map).

The Text Size option can be used to adjust the font size of the graph labels.

Table

This table has one row per sample, showing the average single-cell secretion intensities of each analyte (columns). Under Format Options, you can choose whether the displayed data is background subtracted (Threshold Signal Data on) or raw (Threshold Signal Data off), and whether it's log transformed (Log Transform Signal Data on or off).

Signal Distribution



Grid of Scatterplots

This graph displays a vertical scatterplot of either the secretion intensity of each single cell (y-axis) for each analyte (x-axis), or if samples have been averaged, the average secretion intensity (y-axis) of each analyte (x-axis) by each sample. Under Format Options, you can select Raw values by deselecting Threshold Signal Data or subtract the background by selecting Threshold Signal Data. (Default). You can also select Log Transform Signal Data to display logarithmic values. The option works with either background subtracted or raw values.

Each vertical scatterplot also has a percentage value at the top of the vertical plot for an analyte. The percentage above each analyte specifies the percentage of single cells secreting the corresponding analyte.

You may also select Show Strong Signals, in which case the graph will show you what percentage of analyte readouts are above a specified intensity (1000 RFU by default). If enabled, this is what the percentage above each analyte refers to. Show Strong Signals also adds another percentage for all the analytes in a larger font. This larger percentages at the top of the graph refers to the total percentage of readouts (across all analytes) that is above the specified intensity.

Scatterplot

This graph displays a vertical scatterplot of the secretion intensity of each single cell (y-axis) for each analyte (x-axis). Under Format Options, you can choose whether the displayed data is raw or background subtracted by enabling Threshold Signal Data (Default). You can also select Log Transform Signal Data on (the default) to use Logarithmic values or off. Logarithmic values can be used either with raw or subtracted background options.

Signal Intensity Violin Plot (Single-Cell Signaling)

The Signal Intensity Violin Plot is only present for Single-Cell Signaling Projects. The signal intensity violin plots display the distribution of single cell protein intensities across the selected samples. Each "violin" corresponds to a specific sample or sample group, with the width of the violin indicating the frequency of the intensity specified on the y axis.

The data is log transformed before being displayed. If the Normalize Data button is selected under Filter Selection, the data is also first normalized to the average signal intensity of the housekeeping control protein (Alpha Tubulin).

Under Format Options, The Violin Type option changes how the violin is displayed. The With box plot option adds a box plot to the violin, with the middle horizontal line of the box plot representing the median of the sample, the top and bottom of the box indicating the third and first quartiles, respectively, and the whiskers indicating the maximum and minimum values of the sample. The With jittered dots option overlays dots corresponding to each single cell's protein intensity. The dots are jittered for better visibility.

The sample data is organized in rows and columns, with the first Sort By... option under Format Options corresponding to the factor along the columns, the second corresponding to the factor along the rows, and the third corresponding to the color coding (as seen in the key). By default, proteins are displayed on the x axis, but their placement can be switched to one of the above three locations using the Protein Location option. If switched, the Sort By... options will correspond to the remaining three location (e.g., if Protein Location is set to Rows, then the three Sort By... factors correspond to columns, x-axis and color-coded, respectively).

Note that samples which do not differ in the two selected factors will be automatically averaged together, even if they differ in another selected factor (e.g., if selected factors are Stimulation and Cell Type, samples with the same stimulation and cell type will be automatically averaged even if they are from different donors or donor groups).

Several other options exist for this graph, including Aspect Ratio, to control the overall width to height ratio of the plot and Text Size to adjust the font size of the labels in the graph.

Polyfunctionality / Heterogeneity

Bar Graph

A bar graph showing the polyfunctionality of the selected samples (percentage of single cells secreting two or more proteins). The height of each bar (y-axis) corresponds to the polyfunctionality of the corresponding sample (x-axis). The bars are broken down by the percentage of single cells secreting exactly 2, 3, 4, and 5+ proteins (different shades of orange).

Pie Chart Grid

A pie chart showing the polyfunctionality (for single-cell secretome data) or heterogeneity (Single-Cell Signaling data) of the selected samples (percentage of single cells secreting two or more proteins). The percentage in the middle of the pie chart corresponds to the polyfunctionality/heterogeneity. The pie chart

itself is broken down by the percentage of single cells in each sample (rows) with exactly 0, 1, 2, 3, 4, and 5+ observed proteins.

Table

Shows the polyfunctionality (for single-cell secretome data) or heterogeneity (Single-Cell Signaling data) of the selected samples (percentage of single cells secreting two or more analytes). The next columns show the percentage of single cells in each sample (rows) that are secreting exactly 0, 1, 2, 3, 4, and 5+ proteins.

PSI / PII / FHI

The PSI, PII, or FHI data types are presented for different project types.

- Polyfunctional Strength Index or PSI is presented for Single Cell Secretome projects.
- Polyfunctional Inflammation Index or PII is presented for Single Cell Secretome Inflammation projects.
- Functional Heterogeneity Index or FHI is presented for Single Cell Signaling projects.

Bar Graph

A bar graph showing the PSI/PII/FHI of the selected samples (an index combining polyfunctionality/heterogeneity of the sample and the average intensity of the analytes secreted by polyfunctional single cells). The height of each bar (y-axis) corresponds to the PSI/PII/FHI of the corresponding sample (x-axis). The bars are broken down by the PSI/PII/FHI of each cytokine group (e.g., Effector, Stimulatory, Regulatory, Inflammatory, Chemoattractive).

Bar Graph Grid

A bar graph showing how much each analyte (x-axis) contributes to the overall PSI/PII/FHI of each sample (y-axis). PSI/PII is an index combining polyfunctionality of the sample and the average intensity of the analytes secreted by polyfunctional single cells. The sum of all the columns is equal to the PSI/PII of the sample. As in the Protein Frequency Bar Graph Grid, the color-coding of the analytes corresponds to the functional group they are assigned to, and gray bars correspond to analytes that were not significantly secreted by the sample.

Table

Shows the PSI/PII/FHI of the selected samples (an index combining polyfunctionality/heterogeneity of the sample and the average intensity of the analytes secreted by polyfunctional single cells). The next columns show a breakdown of PSI/PII/FHI per cytokine group (e.g., Effector, Stimulatory, Regulatory, Inflammatory, Chemo attractive). The final columns show how much each protein (columns) contributes to the overall index of each sample (rows). The sum of all these columns is equal to the index of the sample.

Polyfunctional Groups / Heterogeneous Groups

The Polyfunctional Groups data type is available in Single Cell Secretome projects.

The Heterogeneous Groups data type is available in Single Cell Signaling projects.

Heat Map

A heat map comparing the frequency at which various functional and polyfunctional groups are secreted by the samples. The displayed data and options are the same as in the bar graph below. The difference is that instead of one bar per functional group per sample, each column corresponds to a functional group, and each square corresponds to the frequency at which it was secreted by the corresponding sample (row). The shade of orange of a square gets darker as the frequency increases.

Bar Graph

A bar graph comparing the frequency at which various functional and polyfunctional groups are secreted by the samples. The bars correspond to the frequency of secretion of the functional/polyfunctional group specified by the dots below it (the dots specify which analytes are present in the group). The frequency is the percentage of single cells secreting exactly the corresponding functional group in the corresponding sample. The bars are color coded by sample. The functional groups are ordered from left to right in decreasing order of frequency (average across all selected samples).

The functional groups may also be sorted using the Func. Group Sorting drop-down under Format Options. To sort by number of cytokines, use Sort by Polyfunctionality in the drop-down. To sort by the Polyfunctional Strength Index, use Sort by PF index. The Sort by Frequency option means that the most common groups are listed first.

The number at the top of the graph corresponds to either the average frequency of the group, the number of analytes in the group, or the average PSI/PII of the group across all samples (depending on how you are sorting the groups). By default, only the first 25 groups get displayed; this number can also be adjusted under Format Options. You can also specify whether you only want to show groups that contain at least X number of analytes. Lastly, you can display clustered functional groups, which automatically clusters the full set of groups into a smaller set, based on similarity of the analytes in the groups.

PAT PCA / HAT PCA

Polyfunctional Activation Topology PCA chart type is presented for Single Cell Secretome projects.

Heterogeneous Activation Topology PCA chart type is presented for Single Cell Signaling projects.

A plot showing the functional and polyfunctional groups of the selected samples. Each circle/bubble corresponds to a single functional group, with the label next to it, showing you which analytes are in the group. The color-coded dots within each circle represent the frequency of single cells that secreted this group in each sample. The overall color of each group corresponds to the color of the sample that secreted the group with highest frequency. The axes correspond to the principal components of the data, i.e., these variables account for as much of the variability in the data as possible. These variables are linear combinations of the specified analytes; the analytes most strongly present in each component are listed at the ends of the axes. If the value of a principal component is high for a particular group, it is more likely to

contain those analytes. The same options are available under Format Options as for the Polyfunctional Groups and Polyfunctional Heat Map graphs.

Table

Shows a breakdown of the functional and polyfunctional (or for Single-Cell Signaling data, heterogeneous) groups secreted/observed by single cells of each sample. If a single experiment is selected, each row corresponds to a functional/polyfunctional group secreted by the sample (in decreasing frequency). The number of cells, frequency (percentage of the sample), and the exact analytes in the group are shown. If multiple experiments are selected, then each row corresponds to a sample, and each column corresponds to the secretion frequency of a single group (labeled by the analyte(s) in the group).

t-SNE

2D t-SNE

As an alternative to the 3D t-SNE visualization (below), this visualization presents a grid of 2D t-SNE transformed scatterplots. T-SNE (t-Distributed Stochastic Neighbor Embedding) is a non-linear dimensionality reduction method that allows a high dimensional dataset to be visualized in two or three dimensions while still retaining most of the information within the data. The input to the t-SNE algorithm is the raw single cell protein intensity data.

Each row of t-SNE graphs correspond to the single data from a specific sample or sample group, with the same t-SNE transformation applied. Each column corresponds to one of the selected proteins, with the scatterplots in the column corresponding to the intensity levels of that protein. High intensities are red, while low intensities are blue. The rows are sorted according to the Sort By... graph options and labeled according to the Label By... graph options.

By default, the single cell data is log transformed before the t-SNE transformation is applied; this can be adjusted under Format Options. The t-SNE transformation is applied with default hyperparameters as follows: theta = 0.5, perplexity = 50, and iterations = 1000. These may also be adjusted, but the default values should generally give good results.

Several other options exist for the 2D t-SNE plots, including Dot Size to control the size of the dots in the scatterplots, and font size to control the label font sizes.

3D t-SNE

A 3D plot showing the t-SNE transformed single-cell intensities of the selected samples. Each dot corresponds to a single cell. T-SNE (t-Distributed Stochastic Neighbor Embedding) is a non-linear dimensionality reduction method that allows a high dimensional dataset to be visualized in two or three dimensions while still retaining most of the information within the data. Practically, this enables visualization and exploration of key differences across high dimensional datasets, and to better understand differences across various sample groups or conditions. Each dot in the t-SNE scatterplot corresponds to a single cell, color coded by sample, sample group, or a characteristic of the cell's secretion profile (e.g., polyfunctionality).

To generate a 3D t-SNE plot for a selected dataset, click the Generate t-SNE Graph button. Note that due to the complexity of the algorithm, it may take several minutes to generate the graph. Once complete, you will see an interactive and color-coded 3D plot of the t-SNE-transformed cell intensities. Clicking and dragging over the graph will rotate the view. Hovering over specific points with your mouse will display more information for that cell, such as which sample it is from and what cytokines it secreted.

The default hyperparameters used for the t-SNE transformation are as follows: theta = 0.5, perplexity = 50, and iterations = 1000. The single-cell data is also log-transformed by default before applying the t-SNE transformation. These parameters may be changed under Format Options, but the default values should be perfectly satisfactory for visualizing and stratifying single-cell data.

Under Format Options, you can update the color-coding of the t-SNE plot. A t-SNE plot may be color-coded according to Cell Type, Stimulation, Donor ID, Donor Group, Experiment ID, Polyfunctionality (how many cytokines each cell secreted), and Functional Group (what was the dominant functional group secreted by each cell).

If you select a different subset of your data, you will need to click the Update 3D t-SNE button to generate a new t-SNE graph that reflects this new dataset. Clicking the Revert to selected 3D t-SNE filters button will revert the selection filters to the dataset used to generate the currently visible t-SNE plot.

3D t-SNE Sample Similarity (Replicate Samples Only)

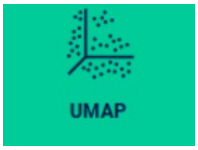
The 3D t-SNE visualization can be used as a method for measuring the similarity of data from replicate samples. When a 3D t-SNE visualization is created with data from two or more replicate samples, a similarity score from 0-100% will be displayed, indicating the overall similarity of the samples' data.

Sample similarity is determined for each pair of samples by looking at the proximity of each data point in sample 1 to a data point in sample 2, and vice versa. The pairwise similarity is averaged across all pairs of samples to obtain the overall similarity of the set of replicate samples. Generally, scores of 80% or higher are expected.

To visually see the similarity between the replicates, select Compare Data Points By below the Visualization Options heading of Format Options and chose Experiment ID. This will assign a different color to each sample and display the graph key in the upper right. In the graph key, clicking individual samples will show/hide those samples in the scatterplot.

Note that the similarity score will only be displayed if replicate samples are used. To use replicate samples, all samples (rows in Project Info) must have the identical annotation values (columns in Project Info). The annotations are what indicate to Data Analysis that the samples are replicates.

UMAP



3D UMAP

A 3D plot showing the UMAP transformed single-cell intensities of the selected samples. Each dot corresponds to a single cell. UMAP (Uniform Manifold Approximation and Projection) is a non-linear dimensionality reduction method (an alternative to t-SNE) that allows a high dimensional dataset to be visualized in two or three dimensions while still retaining most of the information within the data. Functionally, this visualization works the same way as 3D t-SNE, except the data is transformed differently before it's visualized.

To generate a 3D UMAP plot for a selected dataset, click the Generate UMAP Graph button. Note that due to the complexity of the algorithm, it may take several minutes to generate the graph. Once complete, you will see an interactive and color-coded 3D plot of the UMAP-transformed cell intensities. Clicking and dragging over the graph will rotate the view. Hovering over specific points with your mouse will display more information for that cell, such as which sample it is from and what cytokines it secreted.

The default hyperparameters used for the UMAP transformation are as follows: a = 1, b = 1, and iterations = 1000. The single-cell data is also log-transformed by default before applying the UMAP transformation. These parameters may be changed under Format Options, but the default values should be perfectly satisfactory for visualizing and stratifying single-cell data.

Under Format Options, you can update the color-coding of the UMAP plot, by choosing a value in the Compare Data Points By drop-down. A UMAP plot may be color-coded according to Cell Subset, Stimulation, Donor, Donor Group, Experiment ID, Dominant Functional Group (what was the dominant functional group secreted by each cell), and Polyfunctionality (how many cytokines each cell secreted).

If you select a different subset of your data, you will need to click the Update 3D UMAP button to generate a new t-SNE graph that reflects this new dataset. Clicking the Revert to selected 3D UMAP filters button will revert the selection filters to the dataset used to generate the currently visible UMAP plot.

Protein Correlations (Single-Cell Signaling Projects Only)

Chord Diagrams

This visualization displays a grid of chord diagrams, each indicating protein to protein correlations within the corresponding sample or sample group. Two proteins are correlated in a sample if their observed intensities trend in a similar fashion (e.g., if one protein has a higher observed intensity in a single cell, so does the other protein, and same for lower intensities). The correlation is negative if higher intensities of one protein correspond to lower intensities of the other protein.

In each chord diagram, the proteins of the sample or sample group are arranged in a circle, with the chords between proteins indicating their correlation. The higher the absolute correlation (positive or negative), the thicker the chord between the proteins. Positive correlations are increasingly greener, whereas negative correlations are increasingly grayer. The key shows which colors correspond to which correlations.

The grid is organized in rows and columns, with the Sort Sample By option under Format Options giving column headings along the top, and the top/first Then Sort By option giving row labels on the left. Note that samples which do not differ in the two selected factors will be automatically averaged together, even if they differ in another selected factor (e.g., if selected factors are Stimulation and Cell Type, samples with the same stimulation and cell type will be automatically averaged even if they are from different donors or donor groups).

By default, the presented data is log transformed; this may be turned off under Format Options. If the Normalize Data button is selected under Filter Selection, the data is also first normalized to the average signal intensity of the housekeeping control protein (Alpha Tubulin).

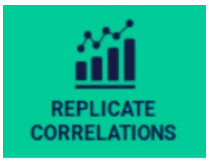
The range of positive and negative correlations displayed in the graphs can be adjusted under Format Options. The Min. and Max. Positive Correlation options control the range of positive correlations displayed (default is all correlations 0-1). Correlations smaller than the minimum will not be shown, and correlations larger than the maximum will be displayed with the same color and chord thickness as the maximum. The Min. and Max. Negative Correlation options function the same way but apply to all negative correlations.

If one of the selected factors has exactly two conditions (e.g., Control and Treated), a third chord diagram will be shown to the right or below the regular two chord diagrams. This chord diagram shows the difference in correlations between the first two plots. Thicker chords indicate greater differences between the samples. The more positive the change (increase in correlations), the bluer the color of chord. Negative changes (decrease in correlations) are shown in red/orange.

The range of displayed differences can be adjusted using the Min. Correlation Difference and Max. Correlation Difference options. By default, the second sample gets subtracted from the first sample. If you wish to flip this (subtract the second sample from the first), select the Flip Two-Group Comparison option. To disable this difference graph entirely, deselect the Enable Two-Group Comparison option.

The label font sizes may be adjusted using the Text Size option.

Replicate Correlations (Replicate Samples Only)



2D Scatterplot

This visualization compares the average protein signal intensity of replicate samples to measure the overall similarity of the data from the replicate samples.

For each pair of samples, the average intensity of each of the selected proteins is compared between sample 1 and sample 2 and plotted as a scatterplot (x-axis is sample 1's average protein intensity, y-axis is sample 2's average protein intensity. Since average intensity of each protein should be close to the same in the two samples, these data points are expected to lie close to the indicated line y = x. An overall R value is computed between the two samples, indicating the overall correlation between the two samples. The R values for all pairs of replicate samples in the dataset are averaged and this overall R value is displayed on the graph.

Note that the similarity score will only be displayed if replicate samples are selected, i.e., all the selected samples have the same annotations.

Chapter 10: CodePlex and Meteor Data Analysis Overview

Since CodePlex and Meteor Data Analysis are similar, the common parts are presented first, followed by separate sections for CodePlex and Meteor.

Once you have processed and obtained data for one or more experiments, you may use the data analysis tools provided in IsoSpeak to begin analyzing the experimental results. You have the option of looking at as many or as few samples as you want. The Data Analysis section can be used to summarize multiple projects runs by importing experiments from other processed projects, so long as the chip types are identical. Meteor experiments cannot be imported into CodePlex projects and CodePlex experiments cannot be imported into Meteor.

Data Analysis Interface

From the Main Menu or the Main Toolbar, click the Data Analysis button. You will see this interface.

The Main Toolbar at the top, which also contains the IsoSpeak logo, lets you go to different program stages, as well as go back to the Main Menu.

You can also use the Export button in the Main Toolbar at any time to export the currently displayed graphs or table. Graphs are output as image files with extension png. while tables can be output either as tab-delimited text files or as .csv files.

The toolbar under the main toolbar contains Select Options on the left and the much wider Select Data Type on the right.

The Select Data Type menu and the associated Select Table/Chart sub-menu under data type let you select your desired visualization (or table), which will show up beneath these toolbars. Selecting a Data Type changes the offered display types on Table/Chart menu. A horizontal scrollbar below the Select Table/Chart sub-menu allows you to scroll left or right to easily access all available options in both Select Data Type and Select Table/Chart toolbars.

On the left, under Select Options, you can select between Filter Selection options and Format Options. Each selection causes its offering options to appear below it.

The Filter Selection options allow you to select specific subsets of the data from the loaded experiment(s). You can filter the data based on stimulations or conditions, samples, donors, donor groups, experiments, markers, and proteins. Adjusting the filters will automatically recreate the displayed graph or table, only showing the samples and sample subsets meeting the criteria of the selected filters. You can also collapse the options that you are not using by clicking on the green triangle to the left of the option headers. Scroll down the list of filter selections using the vertical scrollbar if the selection list is too long to fit on your screen. The offered filter groups are different for Meteor vs CodePlex as listed below.

- Meteor Filters: Condition, Sample ID, Group, Sample Type, Experiments, Sample Index, Proteins
- CodePlex Filters: Stimulation, Donors, Donor Groups, Markers, Experiments, Samples, Proteins

An additional capability of the Filter Selection options is the ability to find average data across different filters (e.g., donors, conditions). For example, if you profiled 5 samples under different conditions and you wish to know the average profile of these samples, simply click the Average button in the Condition filter group. and the displayed table or graph will show the average profile. You can average across any number of criteria at once (for example, average across Conditions and Groups simultaneously). If any of the graphs are averages of multiple samples (see Filter Selection above), error bars will typically be shown by default to illustrate the variability across samples.

Under Filter Selection, General you also have the Quantify Data, to convert signal intensity data (RFU) to concentrations (pg/ml).

Format Options provides various options for fine-tuning the properties, layout, and look of the displayed graphs and tables. For example, you can select how samples are sorted, what gets displayed as the sample title, and various options for Logarithmic scale, Thresholds, Error Bars, Standard Deviation, Min, and Max Graph values.

For any of the displayed graphs, you can use the scroll wheel on your mouse to zoom in an out. Alternatively, you can specify the zoom amount at the top right of the screen.

Some visualizations are exclusive to specific project types (chip families). See Chapter 11 for the visualizations for the CodePlex chip family and Chapter 12 for the Meteor chip family.

Exporting Data

The Export button on the top menu bar exports data to files according to whether Project Info, Data Processing, Data Analysis is being viewed. For Export of Data Analysis, the files exported match to tables and visualizations provided for a chip of a chip family. This section describes export of either a Meteor or CodePlex family chip.

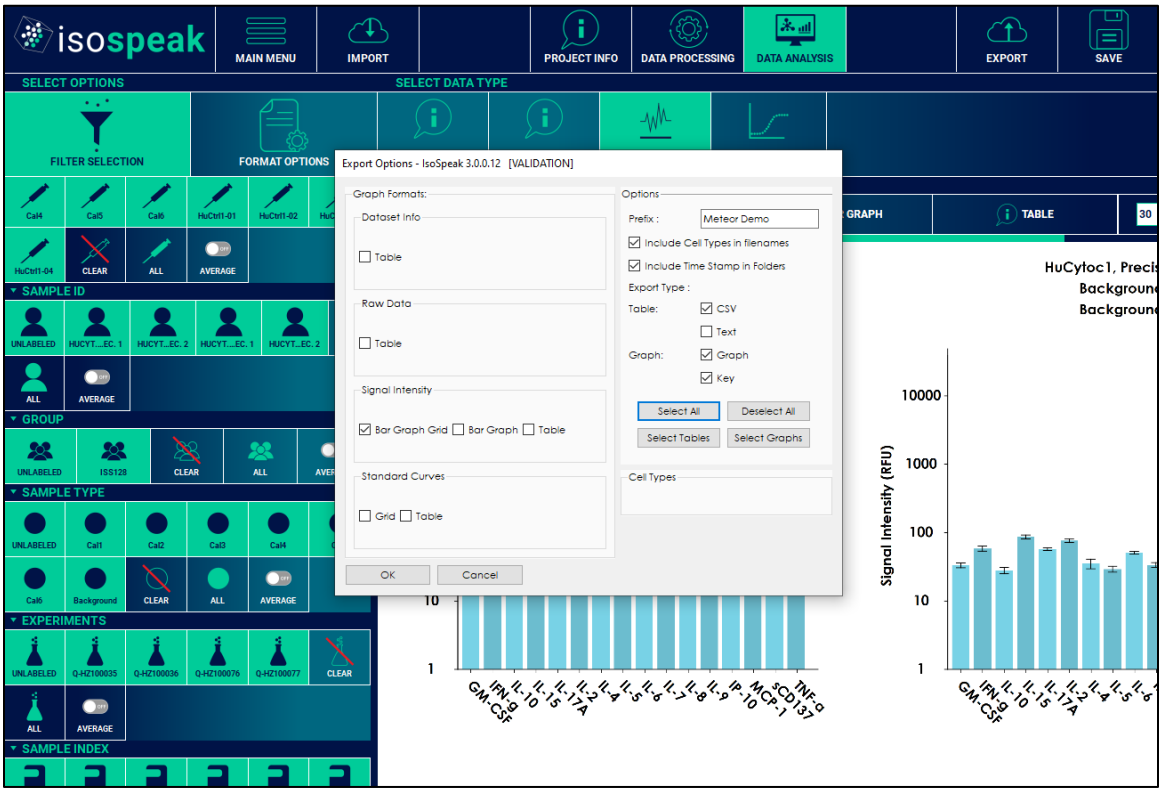
To export any data or visualization from the Data Analysis module, click on the Export button on the topmost menu bar. You will see an Export Options dialog pop up, with the selected option matching the currently visible graph/table.

The Export Options dialog has several options. On the left, under Graph Formats, users can choose from any available table or graph type in data analysis in addition the one being viewed and currently selected. IsoSpeak will export each selected table or graph to files based on your selection in the Options area.

Options allows you to define the prefix of file names and export type of each graph/table, and whether to include cell types in filenames. The default filename prefix is the project name, but you can edit to be any prefix desired within the length limit. Also, by default, cell types are included in the filenames. Under export type, users can choose between "CSV" or "Text" formats or both for tables and between when exporting tables, and "Graph" or "Key" or both when exporting graphs. The graph file exported is a graphical png file. The key file is also a png file but contains the legend that was displayed.

For convenience under Options, Select All (Tables and Graphs), Deselect All (Tables and Graphs), Select Tables (all tables), and Select Graphs (all graphs) will enable or clear the check boxes corresponding.to the selection.

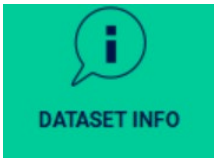
Cell Types box displays the cell samples currently being viewed that will be used in filenames when the Include Cell Types in filenames check box is selected.



Chapter 11: CodePlex Data Analysis Table and Graph Details

This section describes the data types that can be explored in IsoSpeak for a processed CodePlex project, along with the various ways in which each data type can be viewed.

Dataset Info



Table

Displays a table of metadata from Project Info. Each row is a sample row of an experiment with columns Experiment Date, Analyst, Num Samples, Analytes, Donor, Donor Group, Cell Subset, Stimulation. In Project Info, each row is a pair of sample numbers so there are 10 rows per Experiment for 20 samples.

Raw Data

Table

This table displays each sample's signal intensity for each analyte. The first 3 column names and values are chosen from the **Format Options** labels options. First, by "**Label Sample By"** drop-down selection, second by "**Then Label By**" selection and third by the second "**Then Label By"**. The remaining columns are the analyte names and their values. The analyte names can be displayed by hovering over the **Analytes** column under **Project Info** but are also listed under the **Proteins** filter group under **Filter Selection**.

You can also control the type of numerical values. Under **Format Options**, you can choose whether the displayed data is **Raw** (Threshold Signal Data off), a **Threshold Signal** number with background subtracted (Threshold Signal Data on), **Raw and Logarithmic** numbers (Threshold Signal Data off and Log Transform Signal Data on) or **Threshold Signal and Logarithmic** (Threshold Signal Data on and Log Transform Signal Data on).

Under **Filter Selection, General**, there is a **Quantify Data** button that determines whether the signal intensities are displayed as relative fluorescence intensities (RFU) or whether they are converted to pg/ml concentrations using the calibration information for each analyte.

Signal Intensity

Bar Graph Grid

This set of bar graphs shows the average observed intensities (y-axis) in each sample for each of the selected analytes (x-axis). If samples have been averaged into groups, the bar graph will show the average signal intensity (y-axis) of each analyte (x-axis) of the sample group. Under Format Options, you can choose whether the displayed data is Raw, a Threshold Signal (background subtracted), Raw Logarithmic, or Threshold Signal Logarithmic.

Under Filter Selection, General, the Quantify Data button determines whether the signal intensities are displayed as relative fluorescence intensities (RFU) or whether they are converted to pg/ml concentrations using the calibration information for each analyte. If Quantify Data is selected, the units of the graph's y-axis will be pg/ml concentrations; If Quantify Data is deselected, the units of the graph's y-axis will be RFU. Quantified data with background subtraction already applied, regardless of whether Threshold Signal Data is on or off.

Intensities below the limit of detection (LOD) / background threshold of the corresponding analytes will be indicated by fainter bars. Similarly, for intensities above the upper limit of quantitation (ULOQ) the segment of the bar above the ULOQ will be fainter. To display these bars "normally" (without these indications), go to Format Options and deselect Highlight concentrations below LOD or Highlight concentrations above ULOQ, respectively.

Bar Graph

This bar graph is equivalent to the above set of bar graphs, but presents the data in a single graph, with bars color-coded based on the sample. The key to the bar colors is listed under the Bar Graph, below the x-axis and is labeled Key on the left. The bar graphs headings on the Bar Graph Grid will match the key names next to the colors.

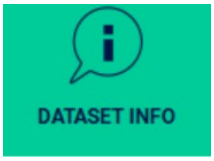
Table

When no sample averaging is enabled, this table will present the same data to the Raw Data table (see above), with each row corresponding to a selected sample and containing the reported signal intensity for each measured analyte. When samples are aggregated (using the Filter Selection → Average buttons), each row will correspond to an aggregate set of samples, and the signal intensities correspond to the average signal intensity across those samples.

Chapter 12: Meteor Data Analysis Table and Graph Details

This section describes the data types that can be explored in IsoSpeak for a processed Meteor project, along with the various ways in which each data type can be viewed.

Dataset Info



Table

Displays a table of metadata from Project Info. Each row is a sample from an experiment (20 samples per experiment) with columns Experiment Date, Analyst, Num Samples, Analytes, Sample ID, Group, Sample Type, and Condition.

Raw Data

Table

This table presents the raw data from the analyzed samples, i.e., per-analyte signal intensities. For each sample, you will see three intra-well replicate values ("triplicates") measured and reported for each analyte, as indicated by the Replicate Index column. By default, these are unitless measurements (RFU, relative fluorescence units). You can select whether the displayed data is background-subtracted or not, using Format Options → Threshold Signal Data. Background levels are computed individually per sample and analyte and each triplicate measurement, using pixels of the signal image adjacent to the signal region. Note: CalB samples are not used to compute background levels – only for calibration curve generation. Under Filter Selection, General, selecting Quantify Data will convert these signal intensities to pg/ml concentrations, if a calibration curve was computed for the sample (using data from the calibrator samples on the same chip).

TABLE

	Experiment ID	Sample Index	Condition	Replicate Index	GM-CSF	IFN-g	IL-10	IL-15
1	Q-HZ100035-230117	12	Cal2	1	5992.00	6110.00	8286.00	7074.00
2	Q-HZ100035-230117	12	Cal2	2	6003.00	6325.00	8864.00	7303.00
3	Q-HZ100035-230117	12	Cal2	3	5674.00	6949.00	9600.00	6672.00
4	Q-HZ100035-230117	14	Cal2	1	4709.00	6120.00	8900.00	7618.00
5	Q-HZ100035-230117	14	Cal2	2	5804.00	6807.00	9757.00	7996.00
6	Q-HZ100035-230117	14	Cal2	3	5486.00	6293.00	9934.00	6970.00

Selecting Format Options → Display Statistics will display statistical information across each sample's triplicate values – the average, standard deviation, standard error and %CV. Format Options → Standard Deviation Type lets you select whether the sample or population standard deviation is reported (and used for the standard error calculation).

TABLE

	Experiment ID	Sample Index	Condition	Replicate Index	GM-CSF	IFN-g	IL-10	IL-15
1	Q-HZ100035-230117	12	Cal2	1	3914.00	3954.00	6152.00	4969.00
2	Q-HZ100035-230117	12	Cal2	2	3894.00	4176.00	6602.00	5224.00
3	Q-HZ100035-230117	12	Cal2	3	3606.00	4713.00	7385.00	4532.00
4				AVERAGE	3804.67	4281.00	6713.00	4908.33
5				STD. DEV.	172.34	390.24	623.95	349.97
6				STANDARD ERROR	99.50	225.31	360.24	202.05
7				% CV	4.53	9.12	9.29	7.13
8	Q-HZ100035-230117	14	Cal2	1	2609.00	4011.00	6459.00	5516.00
9	Q-HZ100035-230117	14	Cal2	2	3729.00	4492.00	7435.00	5742.00
10	Q-HZ100035-230117	14	Cal2	3	3449.00	4192.00	7558.00	4859.00
11				AVERAGE	3262.33	4231.67	7150.67	5372.33
12				STD. DEV.	582.87	242.94	602.15	458.70
13				STANDARD ERROR	336.52	140.26	347.65	264.83
14				% CV	17.87	5.74	8.42	8.54

Note: when data is quantified, you may see ">TOC" in one or more cells of this table. These correspond to reported RFU values which are above the top parameter of the 4PL calibration curve (and thus cannot be accurately quantified).

SELECT TABLE/CHART

TABLE

	Experiment ID	Sample Index	Condition	Replicate Index	IL-6
1	Q-HZ100035-230117	17	Cal1	1	16443.25
2	Q-HZ100035-230117	17	Cal1	2	>TOC
3	Q-HZ100035-230117	17	Cal1	3	13832.18

If for any reason you wish to exclude one or more values from your dataset, you may do so by right-clicking over a data point and clicking Exclude Data Point in the context menu. You can also select multiple values in the table to exclude them all at once. Once excluded, the cells of these values will have a gray background, and the values will not be included in any graphs, statistics, etc. Excluded calibrator data points will also not be included in calibration curve fits. To include a previously excluded data point, select the value, right-click and click the Include Data Point option in the context menu. The cell should no longer have a gray background.

Note: you must first turn off Format Options → Display Statistics to include or exclude data points.

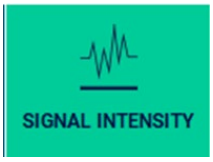
GM-CSF	IFN-g	IL-10
8758.00	13388.00	16327.00
8829.00	13347.00	16796.00
7165.00	13663.00	17955.00

Exclude Data Point

Include Data Point

GM-CSF	IFN-g	IL-10
8758.00	13388.00	16327.00
8829.00	13347.00	16796.00
7165.00	13663.00	17955.00

Signal Intensity

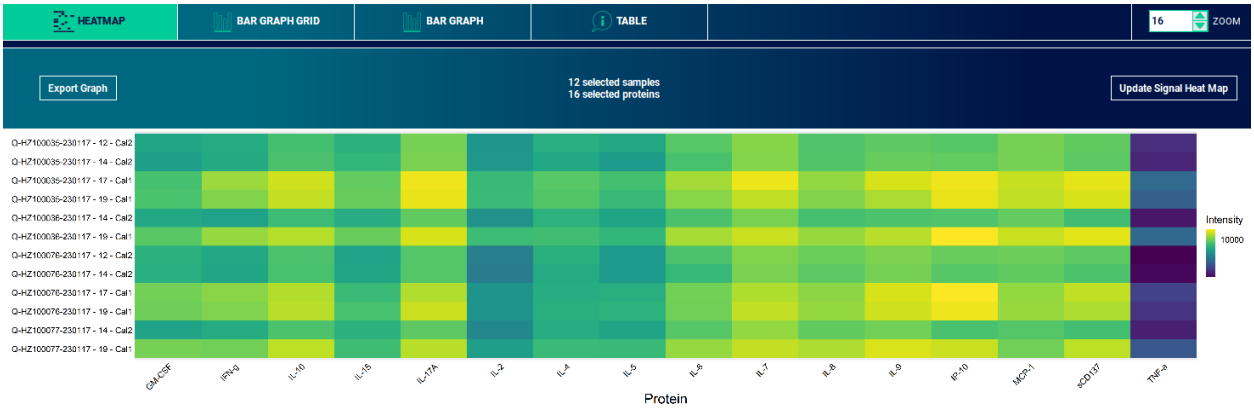


Heatmap

This heat map displays differences in observed protein signal intensities across samples. Each square of the heat map represents the average signal intensity of a given protein in each sample or sample group. Higher signal intensities are yellow, while lower signal intensities are purple.

The sample data is organized in rows and columns, with the first Sort By... option under Format Options corresponding you the factor along the columns, and the second corresponding to the factor along the rows. The sample labels on the left are determined by the Label By... options.

By default, the presented data is log transformed; this may be turned off using Format Options → Use Log Transformed Data. Similarly, the presented data is thresholded but this may be turned off using Format Options → Threshold Signal Data. If the Quantify Data button is selected under Filter Selection, the displayed data is also quantified, using the calibration curve generated for the corresponding chip.

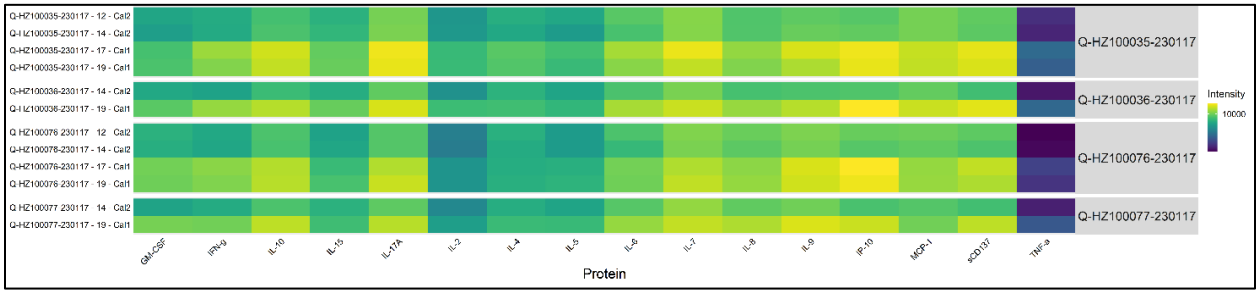


Note: if any values have been manually excluded or, in the case of quantified data, cannot be quantified, they will be represented as gray boxes.

By default, the colors are not scaled (No Scaling option). The colors can further be scaled per column of data in several ways:

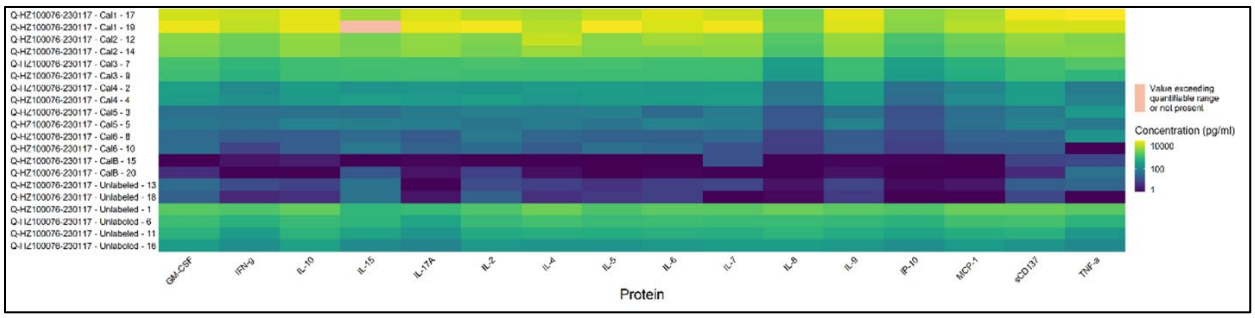
- User-Specified scaling will scale to the values input in the Min. Value and Max. Value options (purple will correspond to the min. value and any value lower than that, while yellow will correspond to the max. value and any value higher than that)
- Standard scaling will divide the data values by the standard deviation of each column of data (Note: will not be applied to quantified data)
- Normalization (0-1) will subtract the minimum data value of the column, and then divide by the range of the data (maximum value – minimum value) in the column (Note: will not be applied to quantified data)

Activating Format Options → Separate Sample Groups will separate and label each set of rows of the heatmap, according to the first Sort By… property (e.g., if you sort by Experiment ID, then each experiment's samples will be separate in the heat map).



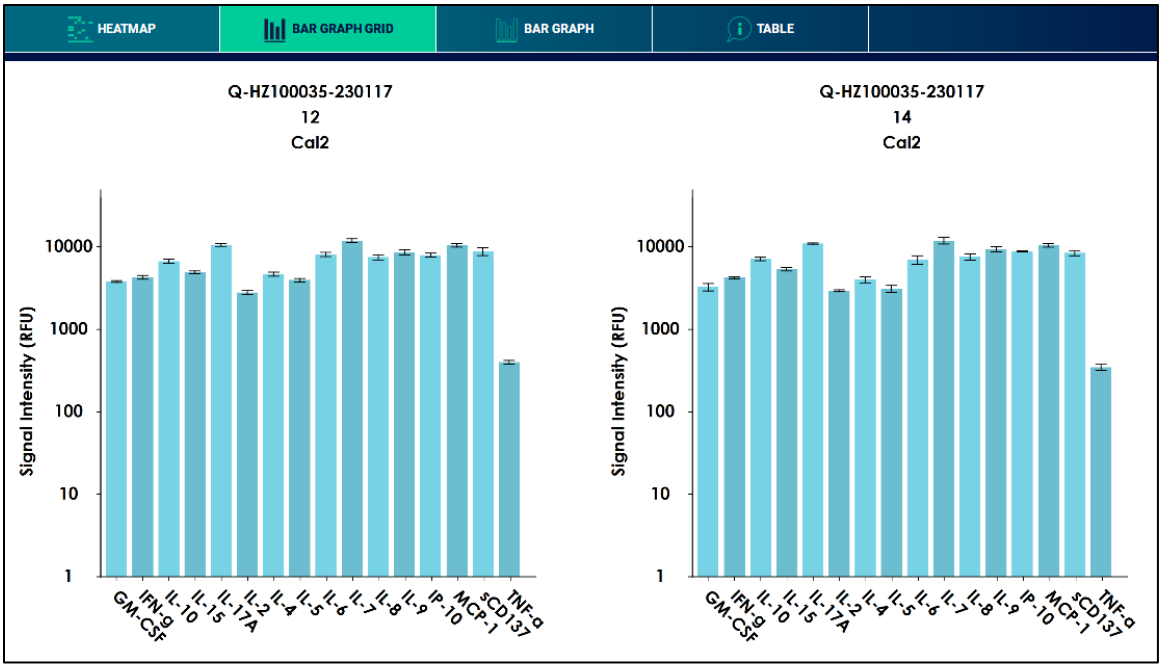
The Text Size option can be used to adjust the font size of the graph labels.

If any data points are not present, either because the underlying raw data was manually excluded or the data could not be quantified, the corresponding box of the heatmap will be highlighted orange:



Bar Graph Grid

This set of bar graphs shows the average observed intensities (y-axis) in each sample for each of the selected analytes (x-axis). If samples have been averaged into groups, the bar graph will show the average signal intensity (y-axis) of each analyte (x-axis) of the sample group.



Under Format Options, you can choose whether the displayed data is raw or thresholded (Threshold Signal Data), or whether the data should be presented on a log scale (Log Scale).

Under Filter Selection, General, the Quantify Data button determines whether the signal intensities are displayed as relative fluorescence intensities (RFU) or whether they are converted to pg/ml concentrations using the calibration information for each analyte. If Quantify Data is selected, the units of the graph's y-axis will be pg/ml concentrations; If Quantify Data is deselected, the units of the graph's y-axis will be RFU.

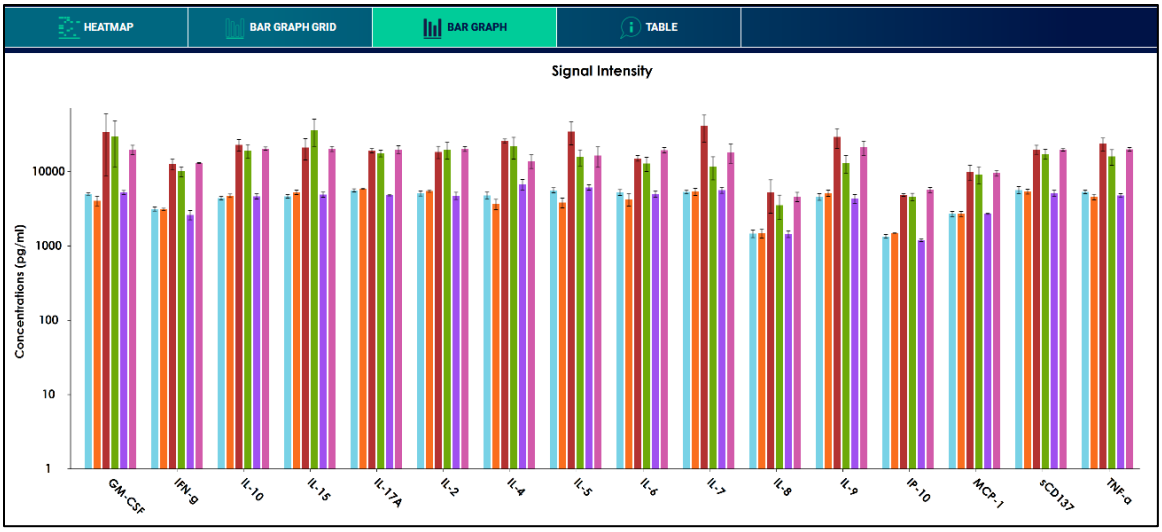
By default, you will see error bars that represent the standard error across the triplicate measurements of each sample – the "intra-sample" standard error. This is the same standard error value that is displayed when you go to the Raw Data table and enable the Display Statistics option.

If any sample averaging is applied, the error bars will represent the standard error across the aggregated samples, i.e., the inter-sample error. This is the same standard error value that is displayed when you go to the Signal Intensity table and enable the Display Statistics option.

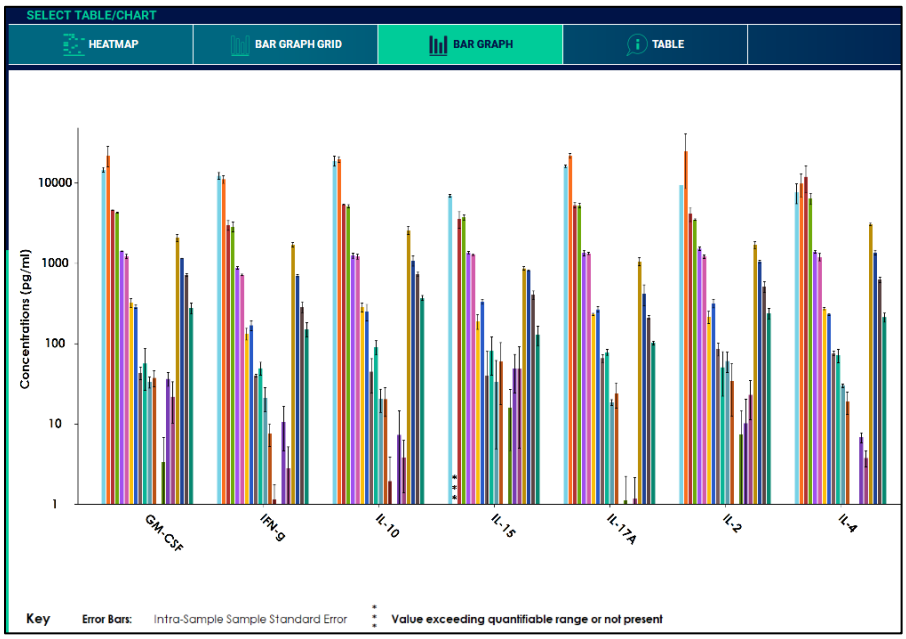
Using Format Options → Error Bar Metric you can change the metric that is represented by the error bar from Standard Error to Standard Deviation. Using Format Options → Standard Deviation Type you can also adjust whether the population or sample standard deviation is used.

Bar Graph

This bar graph is equivalent to the above set of bar graphs, but presents the data in a single graph, with bars color-coded based on the sample.



If any data points are not present, either because the underlying raw data was manually excluded or the data could not be quantified, three vertical asterisks will be displayed in lieu of the corresponding bar in the graph:



Table

This table has one row per sample, showing the average single-cell secretion intensities of each analyte (columns).

Under Filter Selection, General, selecting Quantify Data will convert these signal intensities to pg/ml concentrations, if a calibration curve was computed for the sample (using data from the calibrator samples on the same chip).

You can select whether the displayed data is background-subtracted or not, using Format Options → Threshold Signal Data. If no sample averaging is selected, the intensity values will be equal to the average of the three triplicate values for the sample (as reported in the Raw Data table).

When samples are aggregated (using the Filter Selection → Average buttons), each row will correspond to an aggregate set of samples, and the signal intensities correspond to the average signal intensity across those samples. In this case, selecting Format Options → Display Statistics will also display statistical information across these aggregated samples – the average, standard deviation, standard error and %CV. Format Options → Standard Deviation Type lets you select whether the sample or population standard deviation is reported (and used for the standard error calculation).

HEATMAP | BAR GRAPH GRID | BAR GRAPH | TABLE

	Experiment ID	Sample Index	Condition	GM-CSF	IFN-g	IL-10	IL-15	IL-17A
1	Q-HZ100035-230117	12	Cal2	5889.67	6461.33	8916.67	7016.33	12890.33
2	Q-HZ100035-230117	14	Cal2	5333.00	6406.67	9530.33	7528.00	13401.67
3	Q-HZ100035-230117	7	Cal3	3254.33	3215.67	3834.00	3328.33	5542.00
4	Q-HZ100035-230117	9	Cal3	3496.00	3119.67	4040.00	3401.33	5402.00

HEATMAP | BAR GRAPH GRID | BAR GRAPH | TABLE

	Experiment ID	Sample Index	Condition	GM-CSF	IFN-g	IL-10	IL-15	IL-17A
1	Q-HZ100035-230117	12	Cal2	5889.67	6461.33	8916.67	7016.33	12890.33
2	Q-HZ100035-230117	14	Cal2	5333.00	6406.67	9530.33	7528.00	13401.67
3			AVERAGE	5611.33	6434.00	9223.50	7272.17	13146.00
4			STD. DEV.	393.62	38.66	433.93	361.80	361.57
5			STD. ERR.	278.33	27.33	306.83	255.83	255.67
6			% CV	7.0	0.6	4.7	5.0	2.8
7	Q-HZ100035-230117	7	Cal3	3254.33	3215.67	3834.00	3328.33	5542.00
8	Q-HZ100035-230117	9	Cal3	3496.00	3119.67	4040.00	3401.33	5402.00
9			AVERAGE	3375.17	3167.67	3937.00	3364.83	5472.00
10			STD. DEV.	170.88	67.88	145.66	51.62	98.99
11			STD. ERR.	120.83	48.00	103.00	36.50	70.00
12			% CV	5.1	2.1	3.7	1.5	1.8

If any data points are not present, either because the underlying raw data was manually excluded or the data could not be quantified, the corresponding cell of the table will be gray, with the text "Value >TOC or not present" (when displaying quantified data), or "Value not present" (when displaying RFU data):

	Experiment ID	Sample Type	Sample Index	GM-CSF	IFN-g	IL-10	IL-15	IL-17A
1	Q-HZ100076-230117	Cal1	17	14662.41	12357.98	18960.31	6948.84	16263.04
2	Q-HZ100076-230117	Cal1	19	22215.38	11161.68	19697.12	Value >TOC or not present	21850.82
3	Q-HZ100076-230117	Cal2	12	4607.66	2996.28	5359.84	3563.48	5319.90
4	Q-HZ100076-230117	Cal2	14	4273.01	2864.13	5151.75	3725.40	5282.73
5	Q-HZ100076-230117	Cal3	7	1417.78	877.61	1252.90	1362.09	1348.78
6	Q-HZ100076-230117	Cal3	9	1217.86	719.55	1214.82	1287.93	1334.02

Standard Curves

Meteor Calibration Curves

For each chip, IsoSpeak generates a calibration curve using the data from the calibrator samples on that chip and (optionally) from background calibrator (CalB) samples. As a prerequisite to successfully generating calibration curves, please make sure the proper input information has been provided, as described in Inputting Quantitation Information for Meteor Data.

Four-parameter logistic regression (4PL) curves are fitted to the input data points using the publicly available NPLR curve fitting R package. More information about the package is available here.

The formula of the obtained curves is as follows:

$$y = a + \frac{b - a}{\left(\frac{c}{x}\right)^d + 1}$$

where

- y: RFU response to protein concentration (RFU)
- x: concentration of protein in solution (pg/mL)
- Bottom (a): RFU value of the low asymptote of curve (RFU)
- Top (b): RFU value of the high asymptote of curve (RFU)
- IC50 (c): the x-value of the inflection point of the curve (pg/mL)
- Hillslope (d): the slope or "steepness" of the curve at the inflection point

The inverse of these curves is then used for quantitation, i.e., to obtain concentrations of protein *x* based on RFU response *y*:

$$x = c\left(\frac{b - a}{y - a} - 1\right)^{-1/d}$$

Several curve evaluation metrics (or "goodness of fit" metrics) are automatically computed for each curve:

- Weighted RSSE – The RSSE is the 'residual sum of squares error', it is calculated as the sum of the squared difference between the observed RFU values and the y-value of the regression function. The weighted RSSE is calculated similarly to the RSSE, but each summand is divided by the y2 value of the regression function.
- Weighted Standard Error (SE) – The standard error is the RSSE divided by degrees of freedom to the curve fit, which is defined as (number of calibrator points - number of parameters). The weighted standard error is the weighted RSSE divided by the degrees of freedom to the curve fit, which is defined as (number of calibrator points - number of parameters).
- Weighted R^2 Goodness of Fit – The R^2 goodness of fit is also known as the "coefficient of determination" and is the proportion of the variance in the dependent variable (RFU values)

explained by the independent variable (calibrator concentration). To obtain the weighted R^2 goodness of fit, the weighted RSSE is used in the numerator.

- Median % Prediction Error – We predict the concentration corresponding to each calibrator RFU observation using the curve we fit, and then produce a list of the relative percentage error for each prediction and take the median. Note: if RFU observations for calibrators lie outside of the asymptotes, they are excluded from this calculation since it is impossible to predict the concentration for such points.

In addition, the LLOD (lower limit of detection) and the UL (upper limit) is calculated for each curve:

- LLOD – The LLOD is a "backfit" concentration calculated from an RFU value corresponding to the bottom asymptote of the curve plus two standard deviations of the RFU of the calibrator with the lowest protein concentration (typically either Cal6 or CalB, depending on whether background calibrator samples are included in the curve fit):

$$x_{LLOD} = c\left(\frac{b-a}{(a+2\sigma_{min\ calibrator})-a}-1\right)^{-1/d}$$

- UL – The UL is the concentration of the highest calibrator level where the following equation is true:

$$\frac{y_n}{y_{n+1}} > \frac{d}{2} * \frac{x_n}{x_{n+1}}$$

where n defines the calibrator level (typically 1 through 6).

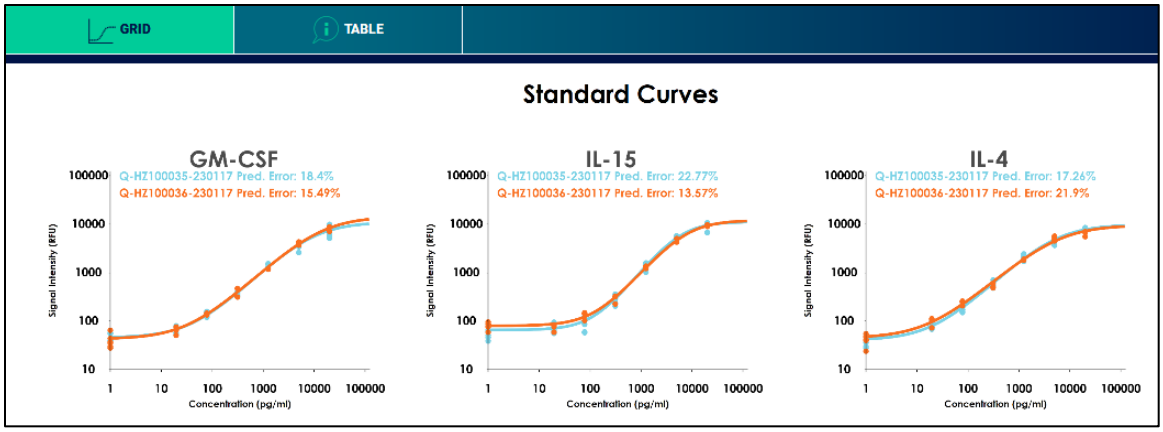
Note that data from samples not labeled as a calibrator (see Inputting Quantitation Information for Meteor Data) will not be used for generating a calibration curve. If any data is manually excluded (see Raw Data Table section), it will also not be factored into calibration curve generation.

In addition, two user-adjustable options exist for calibration curve generation:

- Include Background Calibrators in Curve Fitting – toggle to specify whether Background Calibrator (CalB) data should be factored into the curve fitting.
- Exclude Negative Values from Curve Fitting – in a small percentage of cases, it's possible that a negative value could be reported for a low-concentration calibrator sample (after background subtraction). With this option turned on, those values will not be factored into the curve fitting.

Grid

This set of graphs, one per analyte, displays the calibration curve data for the selected chips: the raw data from each of the calibrator and background calibrator samples, along with the calibrator curve generated from this data. Each chip's data is displayed in a different color. The key displays the first annotation field specified by Format Options → Label By, along with the goodness of fit metric specified by Format Options → Display Goodness of Fit Metric.



The visibility of Background Calibrator data points can be toggled on or off using Format Options → Display Background Calibrator Data Points. Whether or not the y-axis is on a log scale can be toggled using the Log Scale format option, while the minimum and maximum displayed y-values can be adjusted using the Min Graph Value and Max Graph Value format options, respectively.

Table

The standard curves table presents the relevant curve parameters and goodness of fit metrics for all calibration curves generated for the selected samples and analytes. The parameters and curve evaluation metrics defined above reported, along with the number of input data points that are used to generated the curves, and any curve fitting errors returned by the curve fitter.

GRID | TABLE

	Parameter/Metric	GM-CSF	IL-15	IL-4	IL-6	IL-9
Q-HZ100035-230117	Bottom (a)	45.64	65.80	40.68	46.12	57.18
Q-HZ100035-230117	Top (b)	11050.14	11278.72	9822.21	18301.99	33979.64
Q-HZ100035-230117	IC50 (c)	9713.05	5690.99	5262.18	6461.35	14025.61
Q-HZ100035-230117	Hillslope (d)	1.00	1.34	0.99	1.10	0.97
Q-HZ100035-230117	Weighted R^2	0.947	0.920	0.964	0.956	0.949
Q-HZ100035-230117	Weighted RSSE	1.258	1.718	0.969	1.144	1.363
Q-HZ100035-230117	Weighted Standard Error	0.182	0.213	0.160	0.174	0.189
Q-HZ100035-230117	Median % Pred. Error	0.184	0.228	0.173	0.174	0.167
Q-HZ100035-230117	LLOD	17.52	63.02	8.37	12.12	9.17
Q-HZ100035-230117	UL	20000.00	20000.00	20000.00	20000.00	20000.00
Q-HZ100035-230117	# Input Data Points	42	42	42	42	42
Q-HZ100035-230117	Curve Fitting Error(s)	---	---	---	---	---

Chapter 13: Additional Information

IsoSpeak Project File Structure

Creating and saving a project **TestProject** will create a **TestProject.spkpro file**, along with a TestProject folder containing (1) each IsoSpeak experiment file of the project and (2) the corresponding raw data generated by the IsoLight/IsoSpark for each experiment. An example of the full file structure is as follows:

TestProject.spkpro

TestProject\

TestProject_Experiments\

I-IM100001.spk
I-IM100001

I-IM100002.spk
I-IM100002\

I-IM100003.spk
I-IM100003\

TestProject_Images\

ASSAY_I-IM100001\

ASSAY_ I-IM100002\

ASSAY_ I-IM100003\

You can open a project directly through your file browser by double clicking the project (.spkpro) file.

If you move the project file to a different location, please ensure that you also move the associated folder.

Using the **Import** button, raw data from the instrument (one folder per experiment) will be copied to the TestProject_Images folder of a project. You may also copy these files yourself to this folder using File Explorer.

Individual IsoSpeak experiment files (.spk) are stored in TestProject_Experiments.

The default location for saving all IsoSpeak projects is under **\Documents\IsoSpeak Data**. This location can be changed by going to **Options, Set IsoSpeak Data Location…** and choosing a different folder.

Citrix ShareFile for IsoSpeak Updates and Data Quality Analysis

In order to automatically receive updates to IsoSpeak and to allow Bruker to quality control your data, you will receive a **Citrix ShareFile** account, giving you access to an online cloud repository.

In order to facilitate automatic software updates, uploading data for QC to ShareFile, and downloading QC'ed data from ShareFile (see Chapter 6 and Chapter 7 Data Quality Control), we also strongly recommend installing and logging in to the Citrix Files application on your IsoSpeak laptop.

To check if Citrix Files is installed on your laptop and you are logged in with your ShareFile credentials:

Click the ^ button at the bottom-right corner of the Windows Taskbar

Click the Citrix Files icon.

Click the ··· button at the top right of the popup window, and click Settings

Verify that your ShareFile account credentials (name and email address) are visible.

If you do not see the Citrix Files icon in step 2, you likely do not have Citrix Files installed. You may download it at

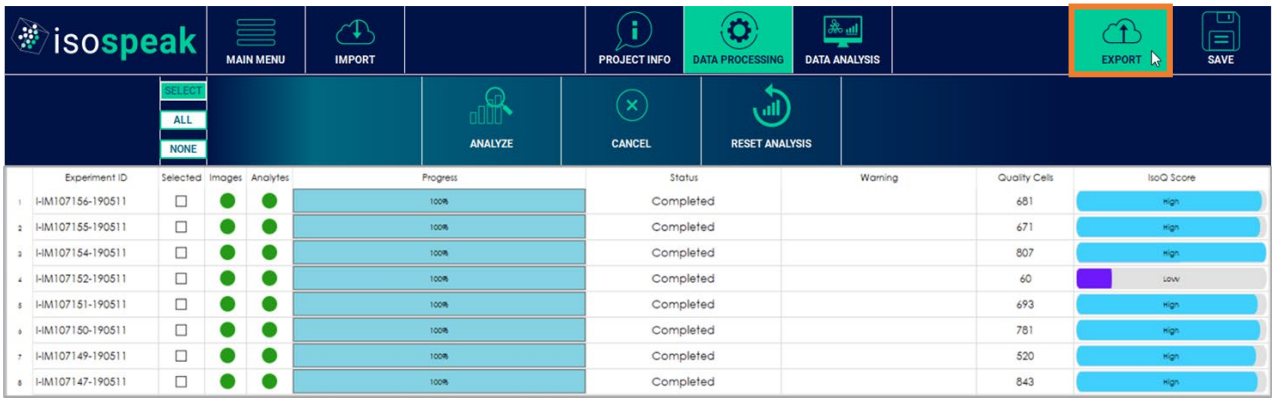
https://isoplexis-sharing.sharefile.com/apps/win. You may need to log in with your ShareFile credentials.

If you do not know your ShareFile account information, or have forgotten it, please contact support@isoplexis.com.

If Citrix Files is installed and you are logged in, IsoSpeak will check for updates on program startup. You will be notified if there are updates available.

You may also check to see if there are updates or obtain the latest version of IsoSpeak by going to S:\Folders\IsoSpeak Updates and running the installer with the highest version number (e.g., IsoSpeak 1.2.3.msi).

Exporting Data from IsoSpeak for Quality Analysis



Once data processing of your samples has completed (see Chapter 6 and Chapter 7), select the experiments you wish to export (e.g., the ones with purple IsoQ scores) and upload to ShareFile. Click the Export button at the top of the screen. In the subsequent popup, select Selected Experiments.

Export your project to a location on your hard drive.

IsoSpeak will create a zip file of your project. The zip file will also contain the instrument logs, if any, associated with the experiments. This may take around 1-2 hours depending on the number of uploaded experiments.

Option 1: Using Citrix ShareFile for Data Quality Analysis

Open your internet browser (e.g., Google Chrome)

Go to https://isoplexis-sharing.sharefile.com/Authentication/Login

Log in with your provided ShareFile account

To upload data, go to **Folders, Shared Folders, [Your-Company-Name], Data Requiring QC**

Click the  button at the top right, and select **Upload**

Click **Browse Files** and locate your zipped project file.

You should see a progress bar appear. Once the progress bar reaches 100%, the data will be uploaded to ShareFile and accessible by Bruker.

Note: Uploading your project to the cloud may take several hours. Please do not turn off your laptop until it is completed.

Bruker will send you an email notification once it has finished quality control of your data. The QC'ed data will be available under **Folders, Shared Folders, [Your-Company-Name], Qced Data and Reports**

Option 2: Using IBM Aspera for Data Quality Analysis

Another method available to you to upload and receive files is by using **IBM Aspera** at http://isoplexis.ibmaspera.com/

IBM Aspera is a **high-performance data transfer service.** The benefit of this service is that it doesn't require usernames and passwords. Upon support request, your customer support representative will send you a one-time link to send or receive data.

Importing QC'ed Data back into IsoSpeak

To update your project with the QC'ed data, go to **Data Processing** and click the **Import** button at the top of the screen.

In the subsequent popup, click **IsoSpeak Project**, and click **Yes** in the next popup.

Select the QC'ed project .zip file using the file browser. IsoSpeak will then update your experiments accordingly.