Demonstration of the Automated Dicentric Chromosome Identifier and Dose Estimator System (ADCI™) in a Cloud-based Online Environment

Ben C. Shirley[a], Eliseos J. Mucaki[b], Joan H.M. Knoll[a,c], and **Peter K. Rogan**[a,d]

[a]CytoGnomix, London, Canada; [b]Schulich School of Medicine and Dentistry, University of Western Ontario, London, Canada; [c]Department of Pathology and Laboratory Medicine, University of Western Ontario, London, Canada; [d]Departments of Biochemistry and Oncology, University of Western Ontario, London, Canada

Email: info@cytognomix.com

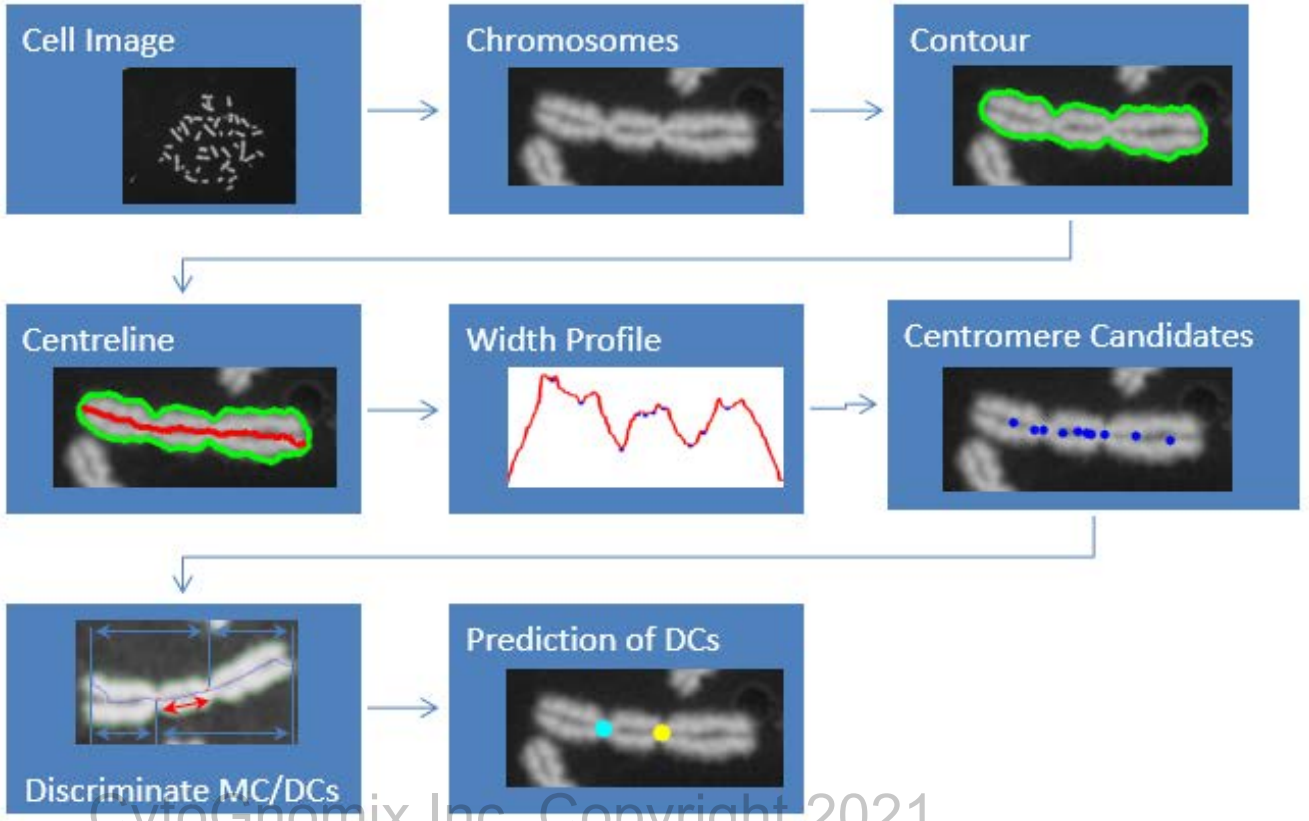
The 3rd Research Coordination Meeting (RCM) of the Coordinated Research Project (CRP) E35010: Applications of Biological Dosimetry Methods in Radiation Oncology, Nuclear Medicine, Diagnostic and Interventional Radiology (MEDBIODOSE)
8 – 12 February 2021

Automated detection of dicentric chromosomes

- The Automated Dicentric Chromosome Identifier and Dose Estimator (ADCI) software completely automates DC detection and estimates biological dose
- ADCI selects appropriate images for analysis, classifies each object as either a centromere-containing chromosome or non-chromosome, further distinguishes chromosomes as monocentric or dicentric using machine learning-based image processing techniques, determines DC frequency within a sample, and estimates biological radiation dose by comparing sample DC frequency with calibration curves computed using calibration samples

- ADCI can process a sample of 500 metaphase images in 3–5 min using a multicore desktop computer system equipped with:
  - Intel i7-6700HQ, 16 GB RAM
  - graphics processing unit: Nvidia® GTX 960M / RTX 2070
- This benchmark estimate is equivalent to ~1.7 images per second, or ~6000 images per hour

- Results fulfill IAEA criteria for triage biodosimetry



ADCI in the Cloud: ADCI_Online

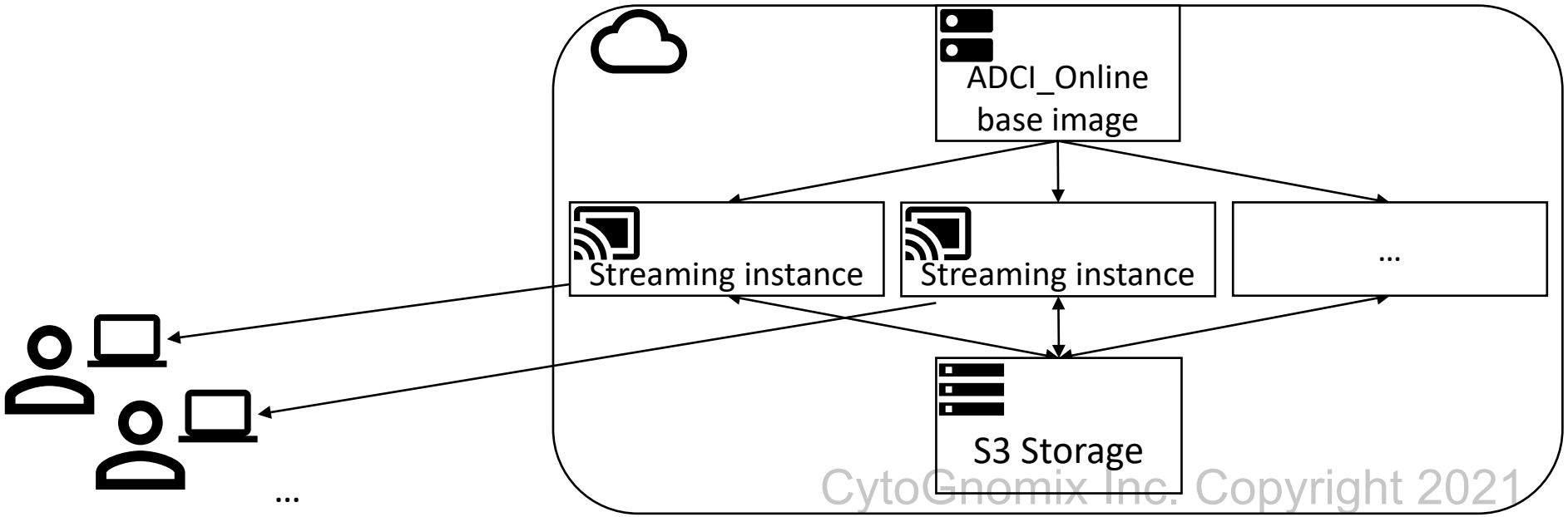
- Unanticipated radiation exposures require rapid dose estimation and discrimination of homogeneous from partial-body exposures.
- Once metaphase cell images are captured by a computerized-microscope system, the same computer typically performs image analysis, during which time the system is unavailable for obtaining images from other samples.
- Outsourcing image analysis to ADCI ™ eliminates this bottleneck and significantly increases overall throughput.
- Accessing ADCI_Online:
  - **Eliminates the need for a dedicated computer system to run ADCI**
  - **Generates the same results as dedicated systems**
  - **Short-term subscription format reduces cost**
  - **Can accommodate on-demand bursts of computing power when necessary**
  - **Can ensure that cytogenetic data is collocated in the same region as the user**
  - **Securely isolates individual user data and protects software analysis from intrusion, disruption and corruption**
  - **Dose estimation can be carried out anywhere there is a reliable internet connection**

ADCI_Online: Differences from MS Windows® version

- Windows®-based ADCI has been ported to ADCI_Online, a secure web-streaming platform (AppStream 2.0, *Amazon Web Services*) that can be accessed worldwide.
- Operationally, ADCI_Online is indistinguishable from the MS Windows version that runs on a dedicated, standalone computer.
- Because the base hardware configuration of AWS ADCI_Online is a 2 processor CPU, 3.75Gb RAM, sample processing is *~3 fold slower* than the recommended Desktop computer running Windows® ADCI (Intel I7-7th gen 4 proc. CPU, 16 Gb RAM, w/o GPU).
- ADCI_Online runs exclusively on the cloud-based system and pixels are streamed to the user through a web browser. The experience is similar to watching a movie in a web browser in that computing resources and disk space are not consumed on the local system. Local keyboard and mouse commands are sent to the cloud-based system to control ADCI_Online.
- ADCI_Online is accessed by streaming. The total duration of each subscribed streaming session can be up to **96 hours** or 4 days (actual clock time).

Scalability

- Although each system running ADCI_Online has less computing power than a high-performance system running ADCI Desktop, the cloud-based nature of ADCI_Online allows for rapid expansion of resources.
- If many samples need to be processed quickly, more computing power can be added (but incurs additional costs). ADCI_Online can be reconfigured quickly within ~15 minutes to:
  - Clone the system as many times as necessary, providing an array of cloud-based systems available for use
  - Increase the computing power of each system running ADCI_Online
  - Both of the above
- A similar array of systems could be achieved by utilizing multiple physical computers, however the ability to expand and reduce available resources as necessary provides flexibility and reduces costs.



Data and Program Security

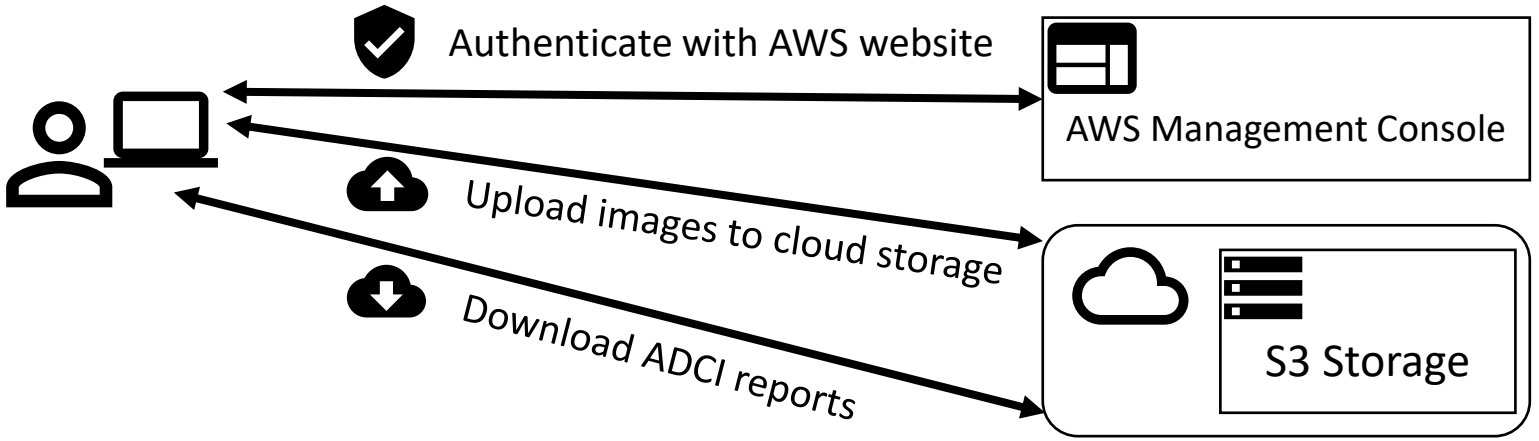
- When a new user accesses ADCI_Online:
  - A UserName is created from their email address
  - An Amazon Web Services (AWS) S3 Bucket is used to store their metaphase image files and reports generated by analysis of the data.
- The user-specific directory in the S3 Bucket is mounted to the cloud-based system, granting the user access to their uploaded images in ADCI_Online.
- Files are encrypted in transit to/from the Bucket (HTTPS protocol) and server-side encryption is applied to all files in the Bucket (AWS Key Management Service)
- File types (e.g. images) are verified by the system upon uploading. Uploaded files are prevented from running as executables on AWS AppStream.
- Internal elements of ADCI software are invisible/inaccessible to the user
- Internet and browser access from within ADCI_Online itself is disabled
- Only files created by ADCI_Online can be downloaded by the UserName that generated them.

Validation of ADCI_Online

- Dose estimates obtained through ADCI_Online were validated against those obtained through ADCI Desktop
- Metaphase images associated with a test sample (HS01) obtained from Health Canada (HC) were uploaded to ADCI_Online and processed. The number of DCs detected matched ADCI Desktop results.
- Other previously processed HC and PHE samples were uploaded to ADCI_Online
  - HC: Dose estimates for homogeneously irradiated samples were generated after application of 4 different image selection models (A_B, A_C, A_D, Automated178981) and matched those presented in *Rad. Prot. Dosimetry* **186(1)**: 42-47, 2019.
  - PHE: Estimates of partial-body dose and fraction of blood irradiated were generated for partially irradiated samples (PHE_E, PHE_F, PHE_G) and agreed with those presented in *Int J Rad Biol*. **96(11)**: 1492-1503, 2020.

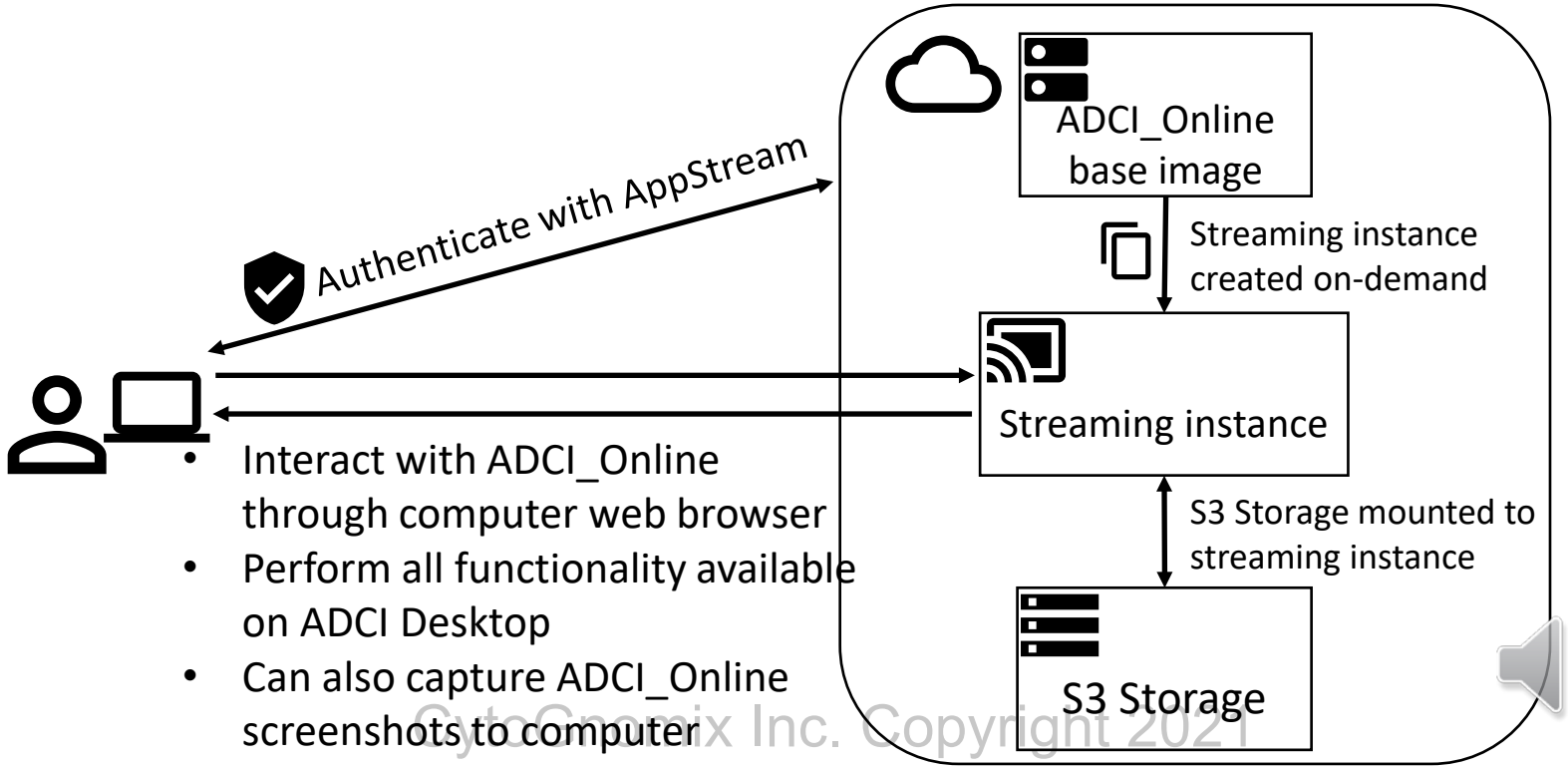
Demonstration: Overview of user interaction

1) Before a subscription period begins, a new user can sign into the AWS Management Console (sign-in credentials provided) and upload metaphase images to cloud storage. This mechanism is also used to download ADCI reports after they have been generated.



2) To access ADCI_Online, the user signs into AppStream in their web browser (sign-in credentials provided) and requests a new streaming session. Behind the scenes, a new streaming instance is cloned from a snapshot of the ADCI_Online system.

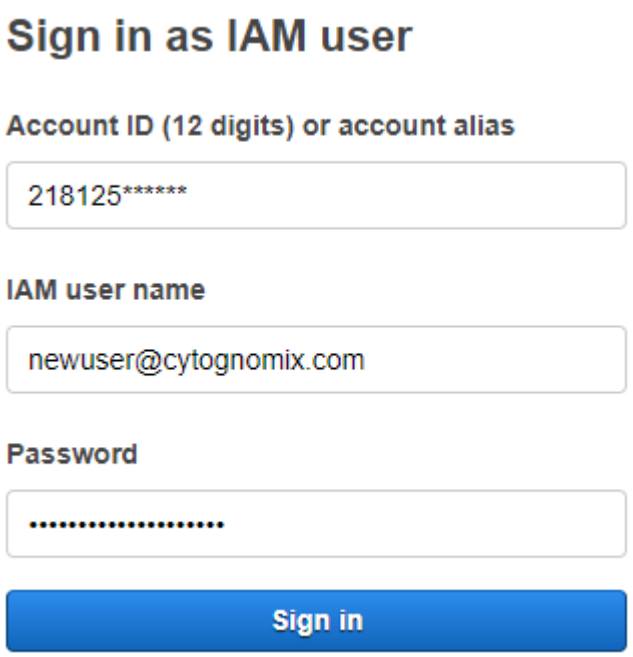
Simultaneously, the user's S3 storage directory is mounted to the streaming instance, allowing the user to access their uploaded metaphase images and save results generated while executing ADCI.

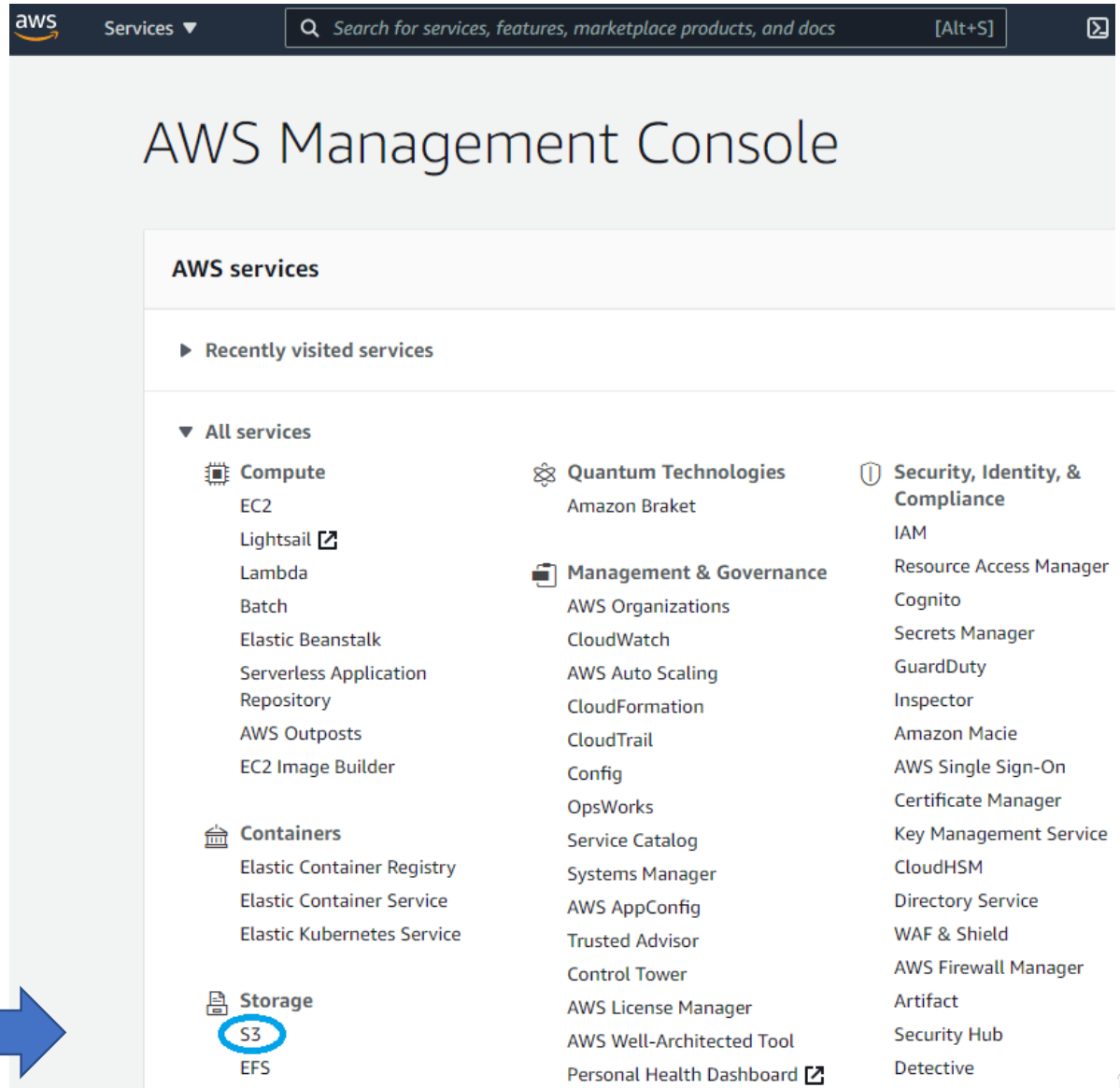


- Interact with ADCI_Online through computer web browser
- Perform all functionality available on ADCI Desktop
- Can also capture ADCI_Online screenshots to computer

Demonstration: Metaphase image upload

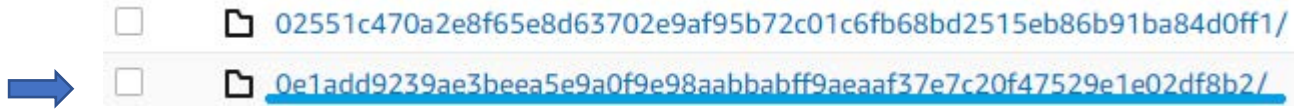
A custom script adds the new user to the ADCI_Online system, configures their S3 storage directory, and *generates a random password for the user*. *An e-mail is sent to the user containing their password*, and the SHA-256 hash of their UserName. The user signs in to AWS and accesses S3 to begin uploading metaphase images to their own directory.

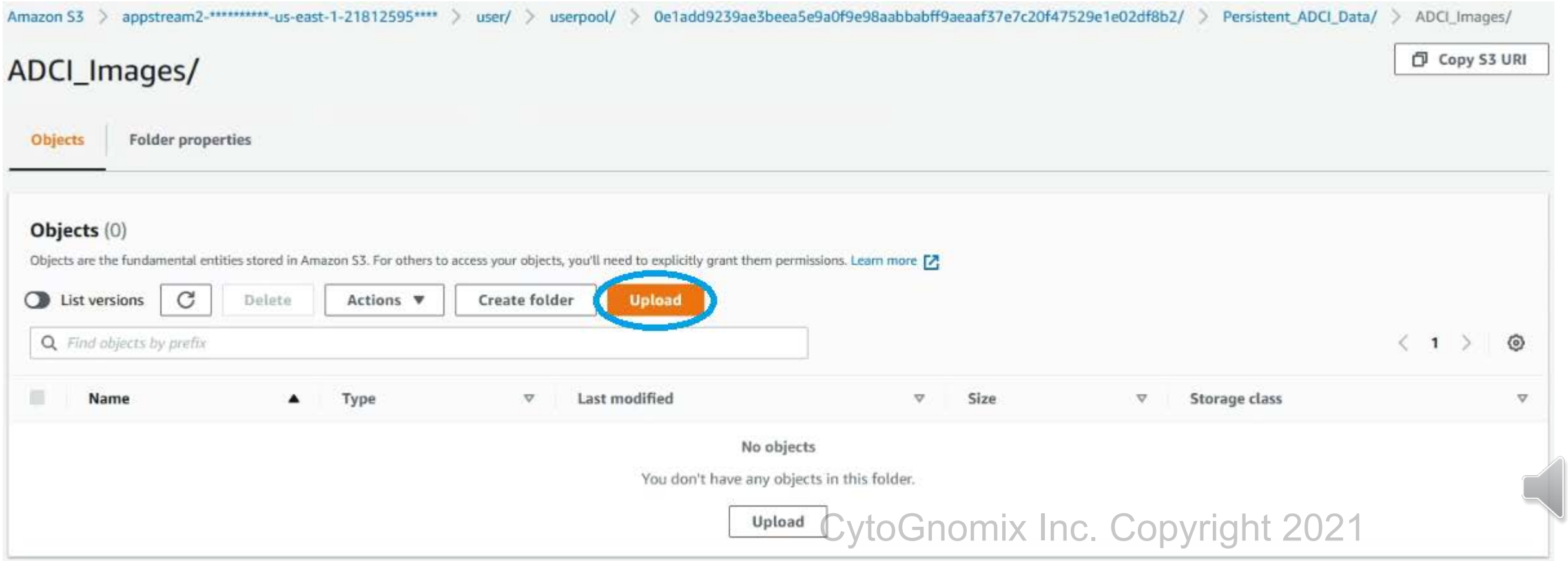




Demonstration: Metaphase image upload (cont.)

Other user directories are present in the user/userpool/ directory of the Bucket, however each user can view the content of their own directory only. Data stored within the "Persistent_ADCI_Data" directory is preserved between streaming sessions. Within that directory, an empty "ADCI_Images" directory awaits metaphase image uploads. Metaphase images for each sample are uploaded into separate folders on the system.



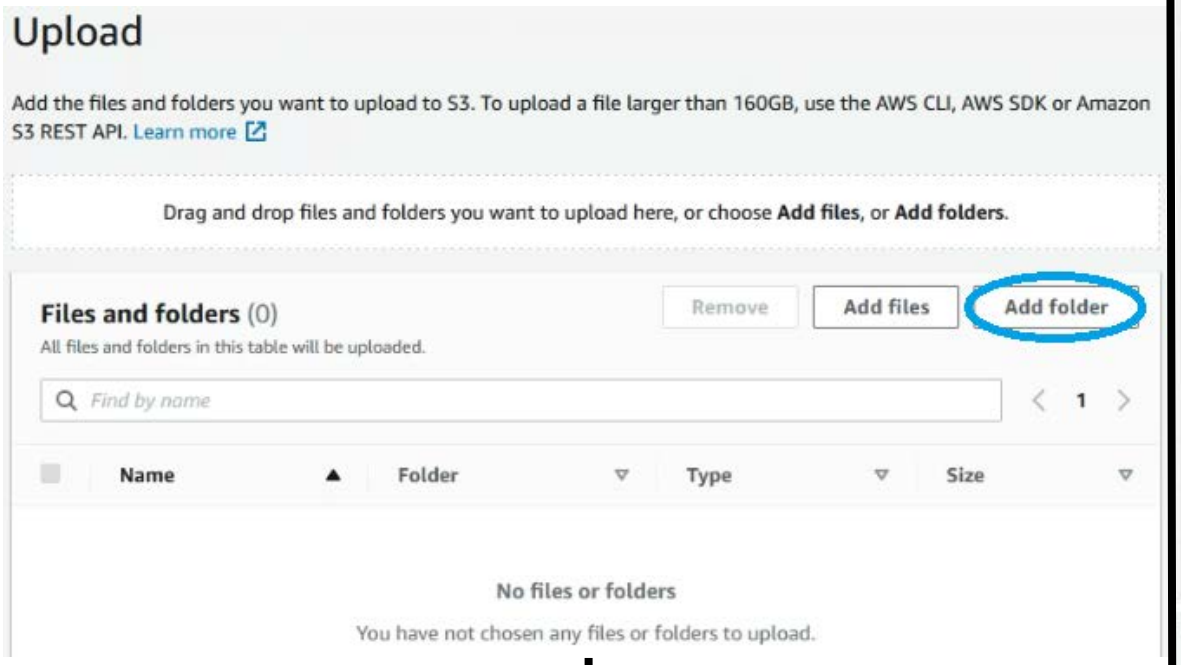


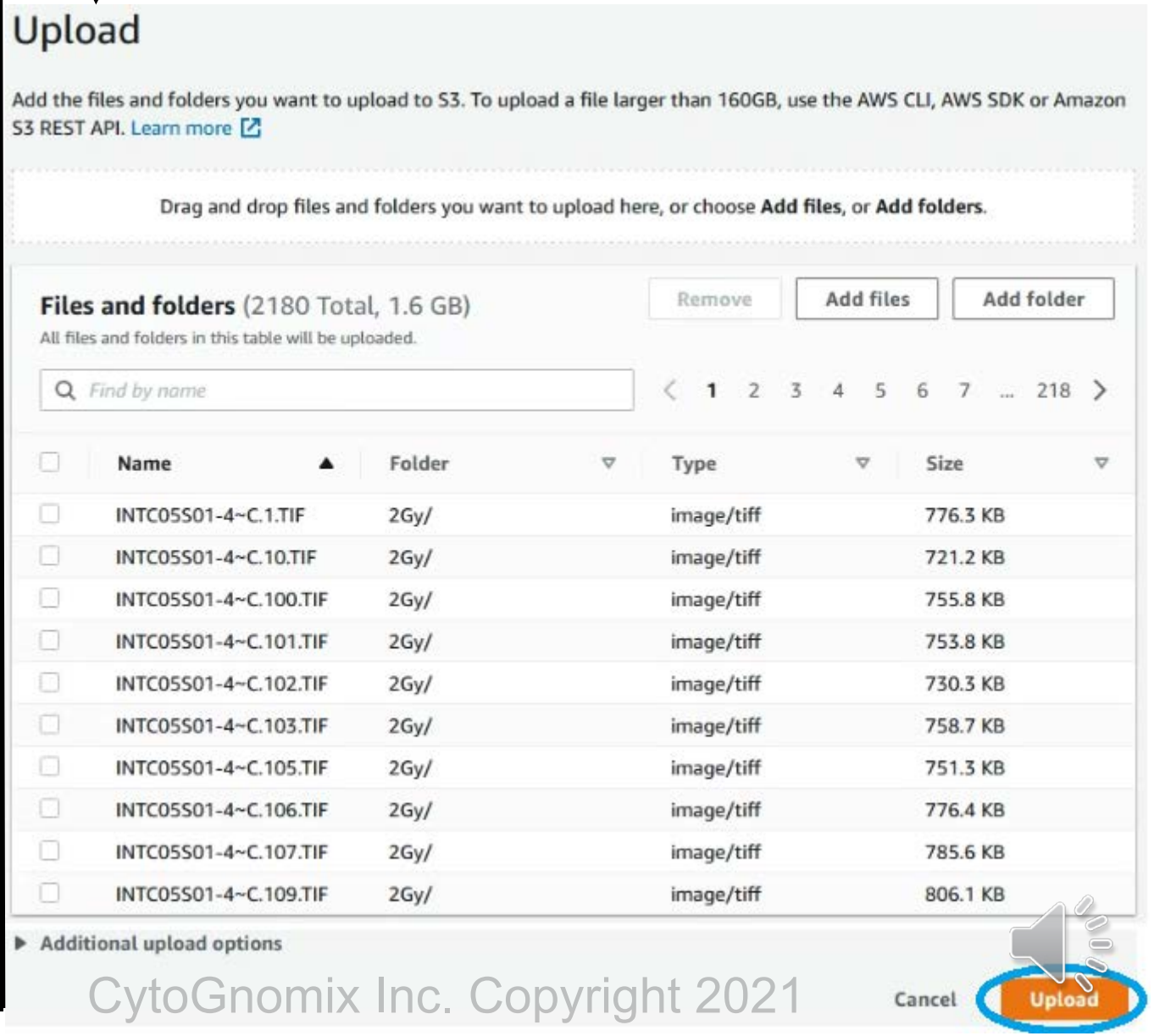
Demonstration: Metaphase image upload (cont.)

Folders of metaphase images can be uploaded to the system.When the user uploads an entire folder to S3 using the "Add Folder" button (circled on the slide), the name of the uploaded folder is the same as the name of the folder on the user's local computer.

ADCI expects each uploaded sample to be stored within its own directory. **In this example, a 2Gy calibration sample is being uploaded to "ADCI_Images/2Gy/"**

Regardless of the folder name on S3, the user provides a sample ID when they create a new sample in ADCI_Online. This ID is used to refer to the sample in all ADCI_Online functions.

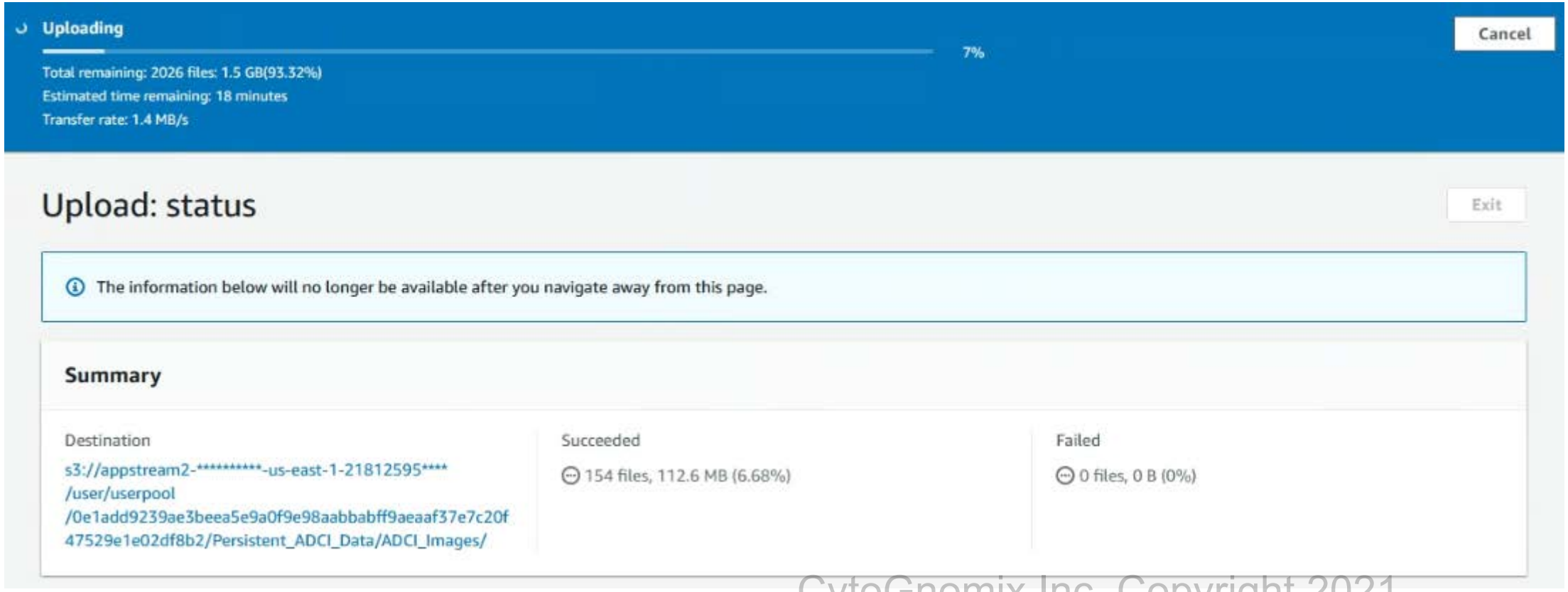




Demonstration: Metaphase image upload (cont.)

Progress of the upload can be seen in the banner at the top of the page. Users can ensure all images have been uploaded successfully by observing the “Succeeded” and “Failed” sections.

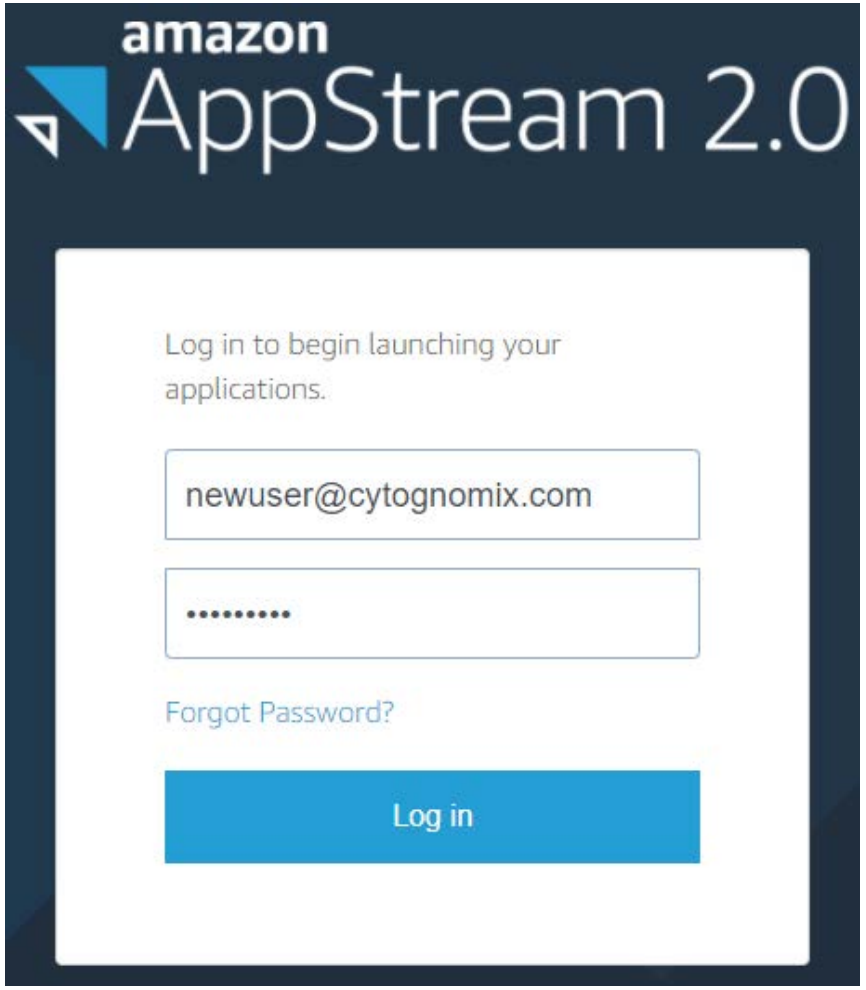
In this example, 2180 metaphase images (1.77Gb) were uploaded to the S3 Bucket in 17 minutes 33 seconds. This is a rate of ~2 images / second.

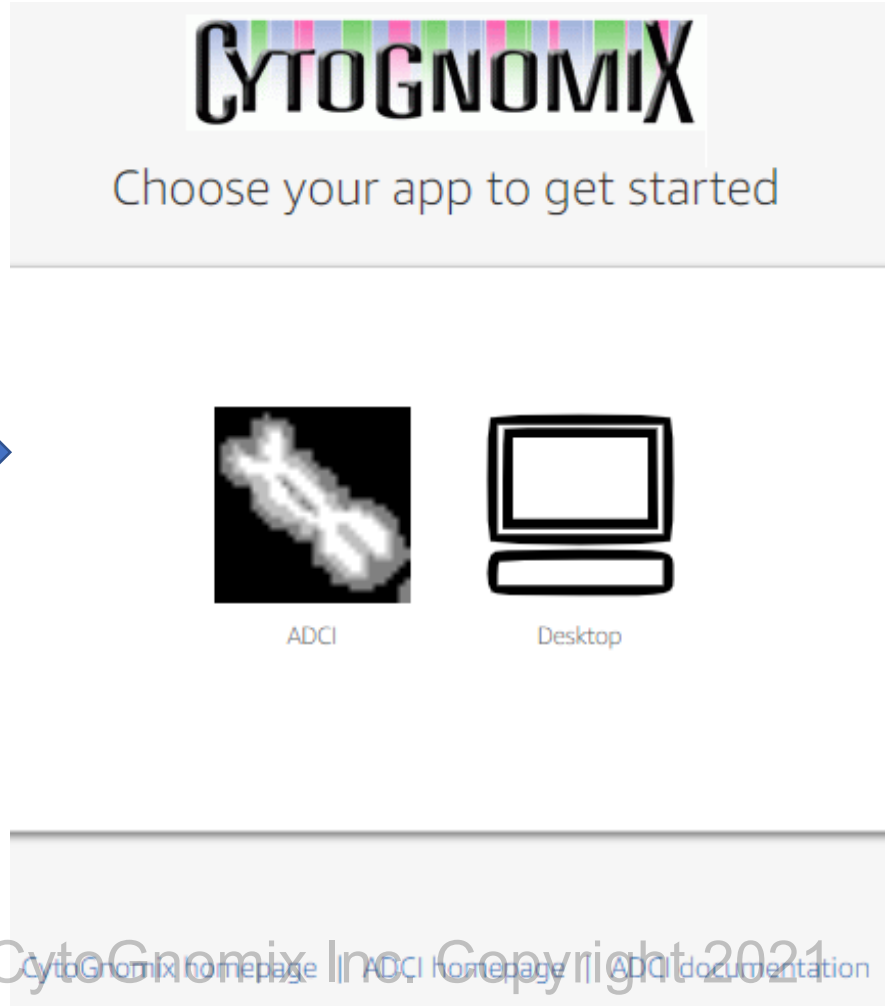


Demonstration: Begin ADCI_Online streaming session

Users receive an e-mail from AWS with a link to *a log-on webpage containing a temporary password and it prompts them to login and update their password*. Once signed in, ADCI_Online can be accessed in an "Application" view with no Windows desktop, or as a "Desktop" view which resembles a remote desktop session.

Users can access ADCI_Online in a web browser or using desktop software provided by AWS.

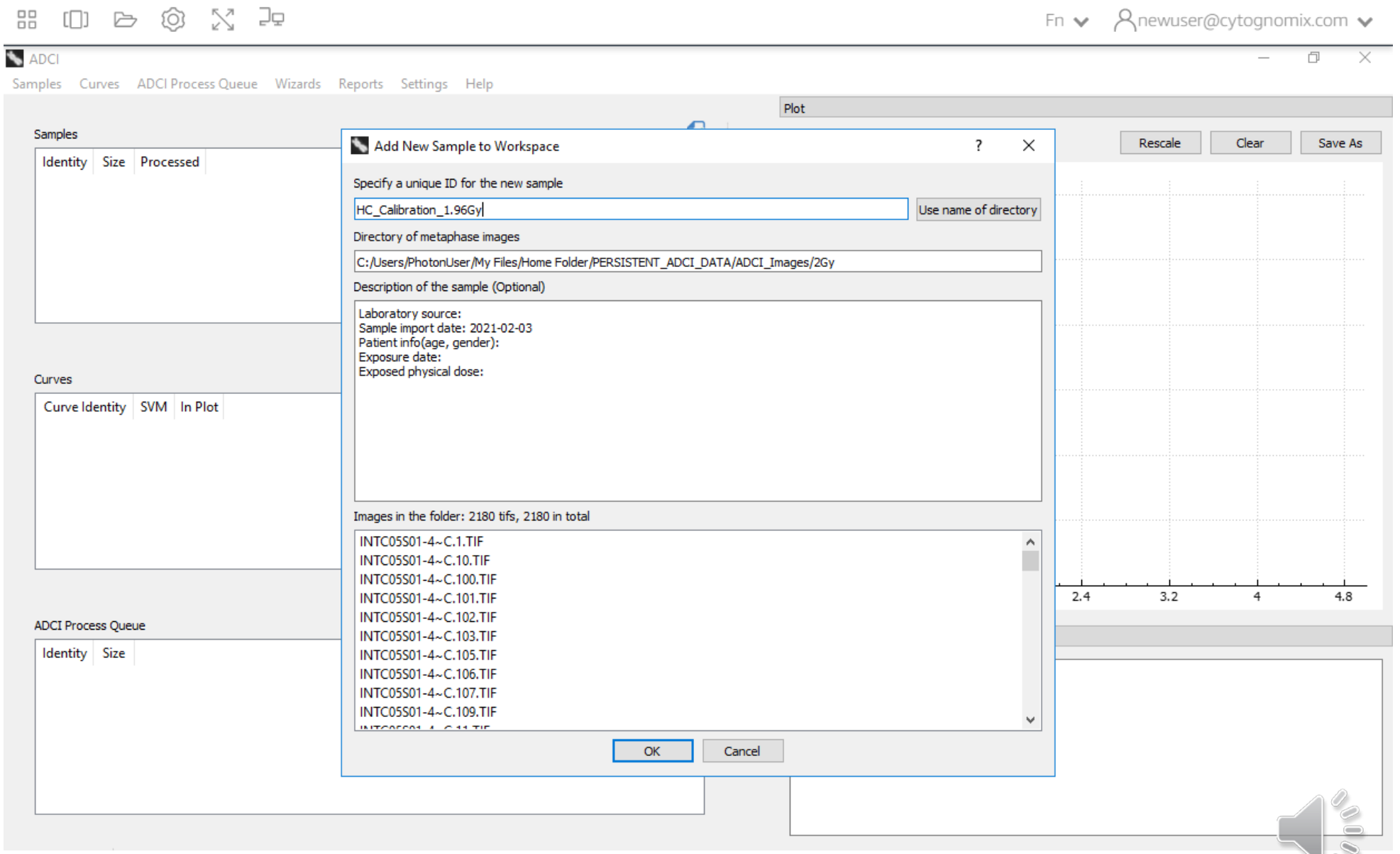




Demonstration: Create a new sample

The content presented here is streamed to the user's web browser.

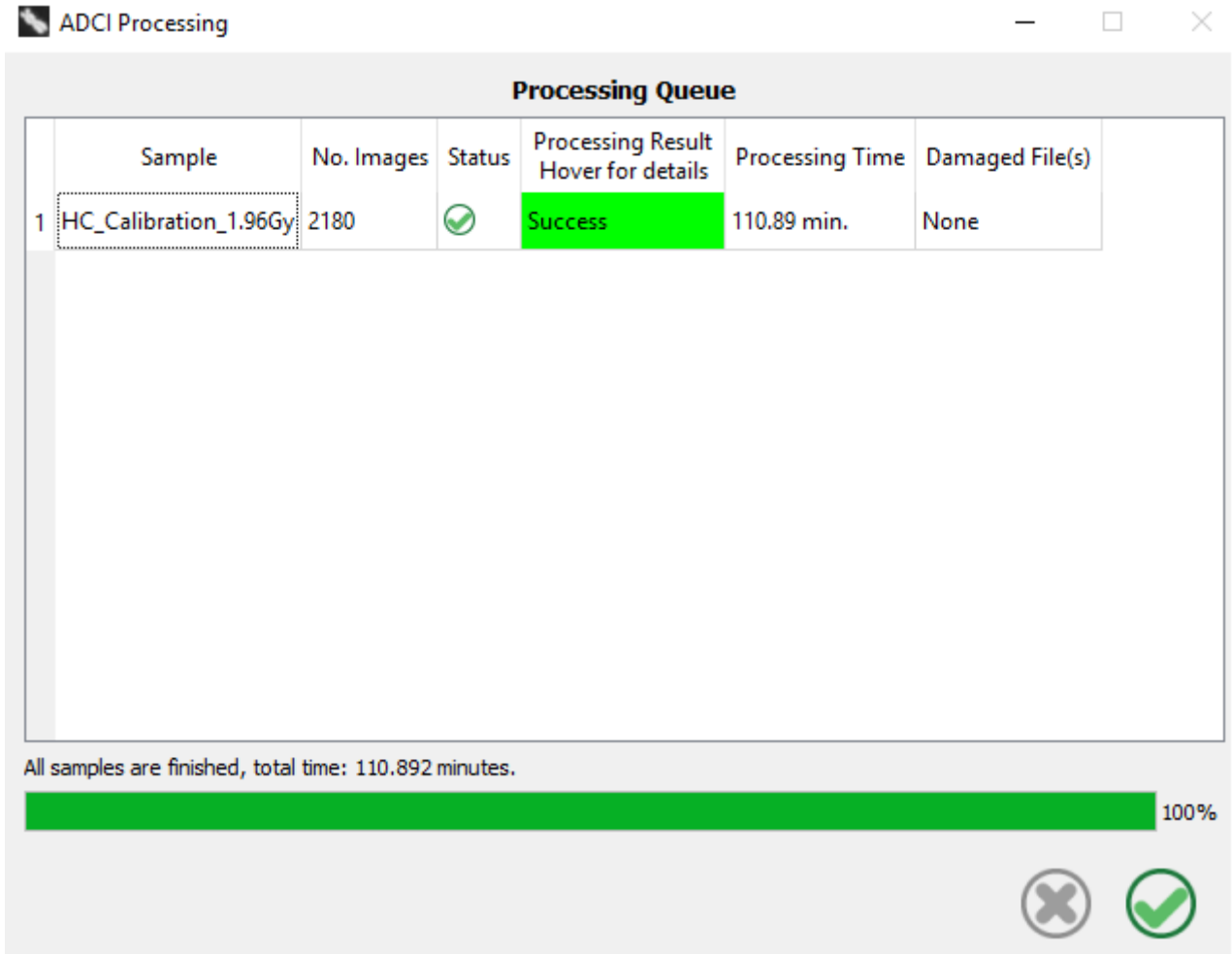
The dose of a calibration sample can be embedded in its sample ID allowing it to be recognized by ADCI automatically. An embedded dose must be in the format of #Gy. For example, the 1.96Gy sample here can contain the text '1.96Gy' or '1.96 GY' allowing ADCI to pre-populate dose fields which generating a calibration curve.

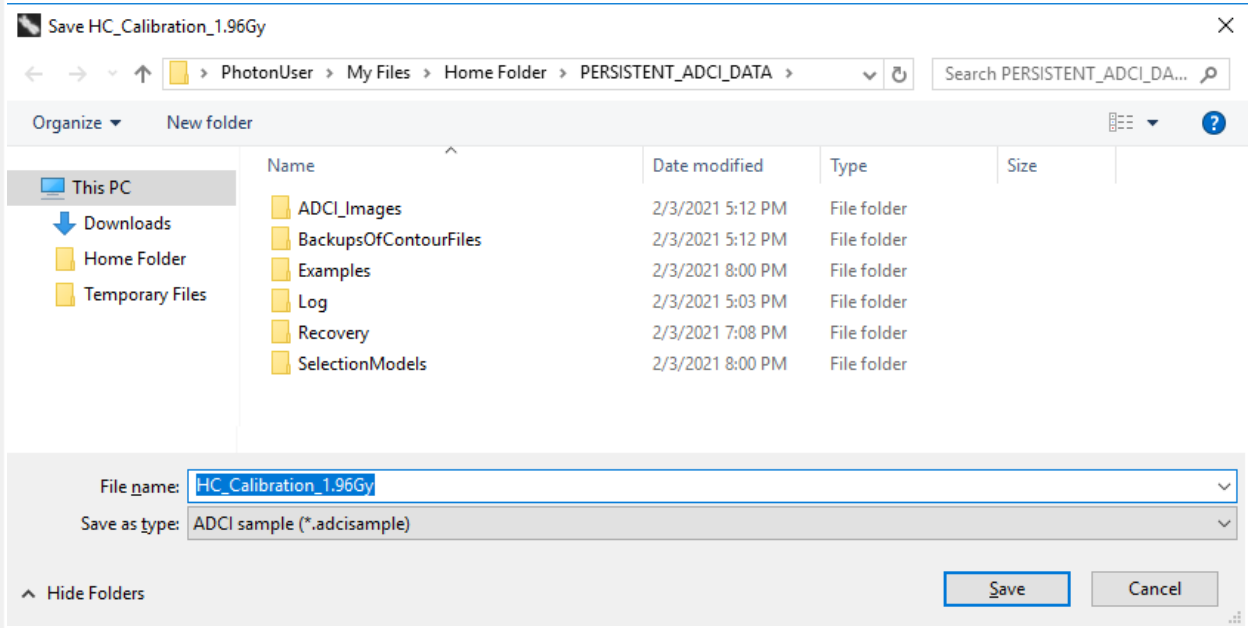


Demonstration: Process a sample

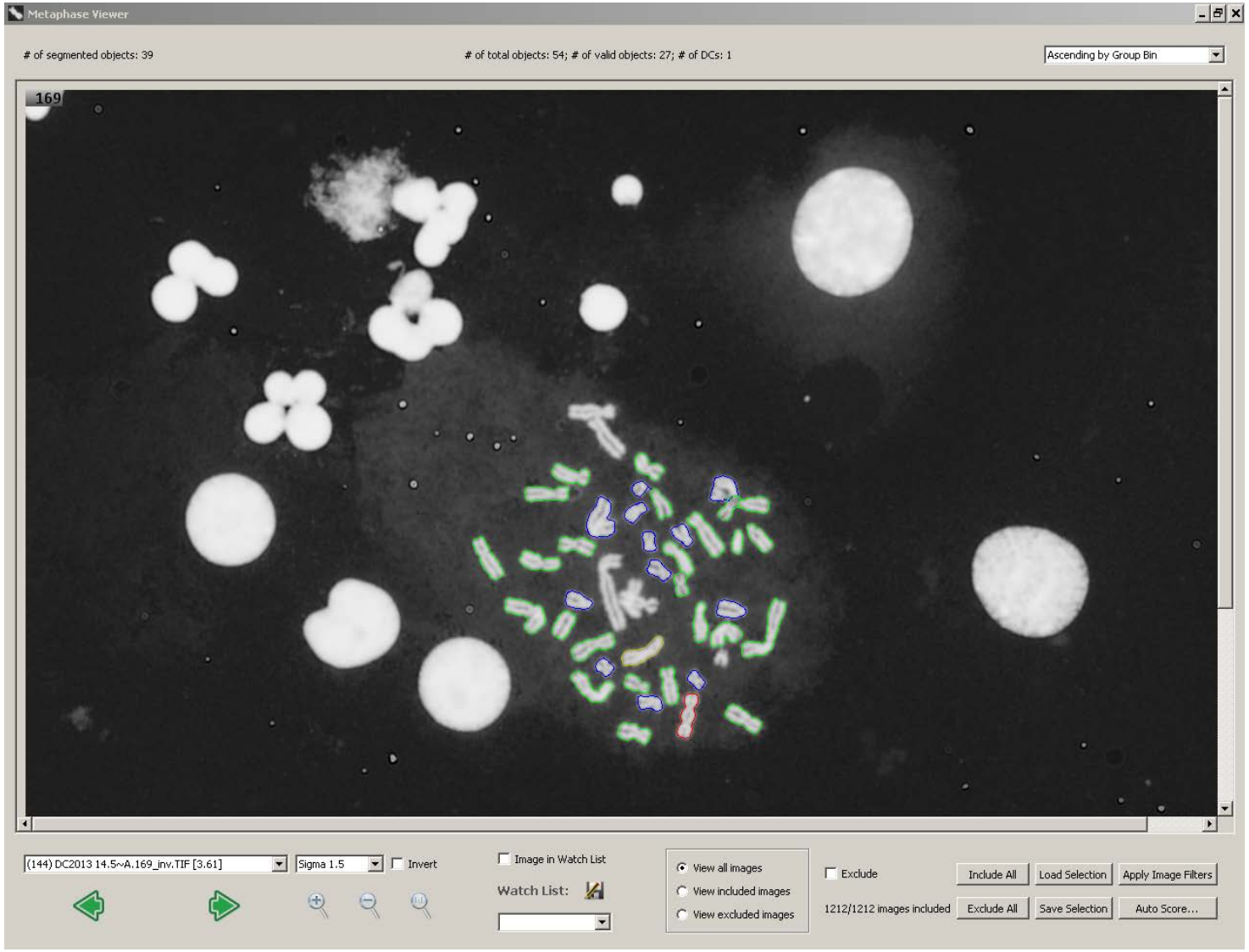
In this example, ADCI_Online processed a 2180 image calibration sample in 110.89 min. This is a rate of ~20 images / min.

When saving samples, curves, or reports ADCI_Online automatically navigates to the “Persistent_ADCI_Data” directory in the user’s “Home Folder”. All files present in the Home Folder are preserved between streaming sessions.





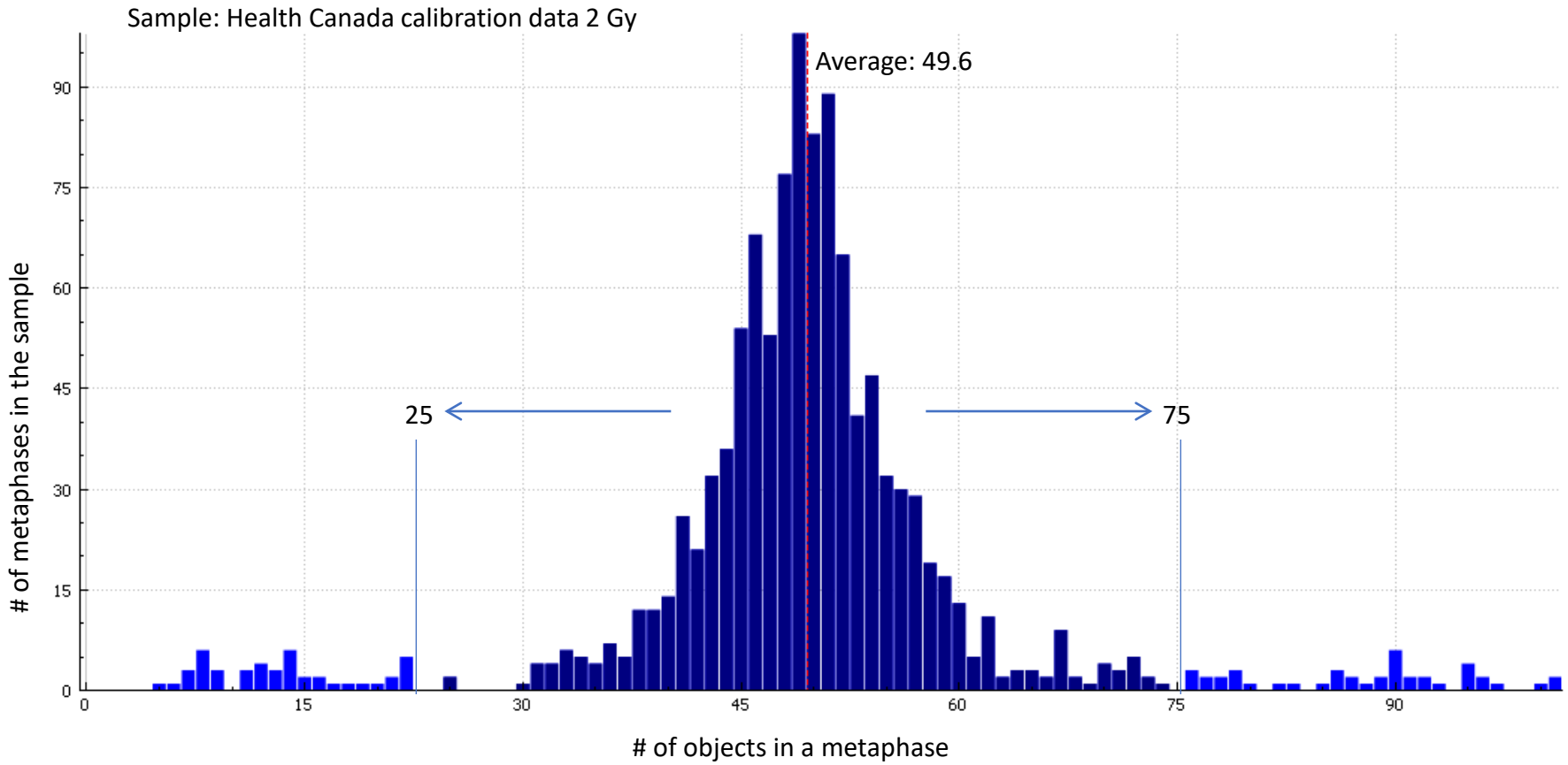
Post-processing: Metaphase cell image viewer



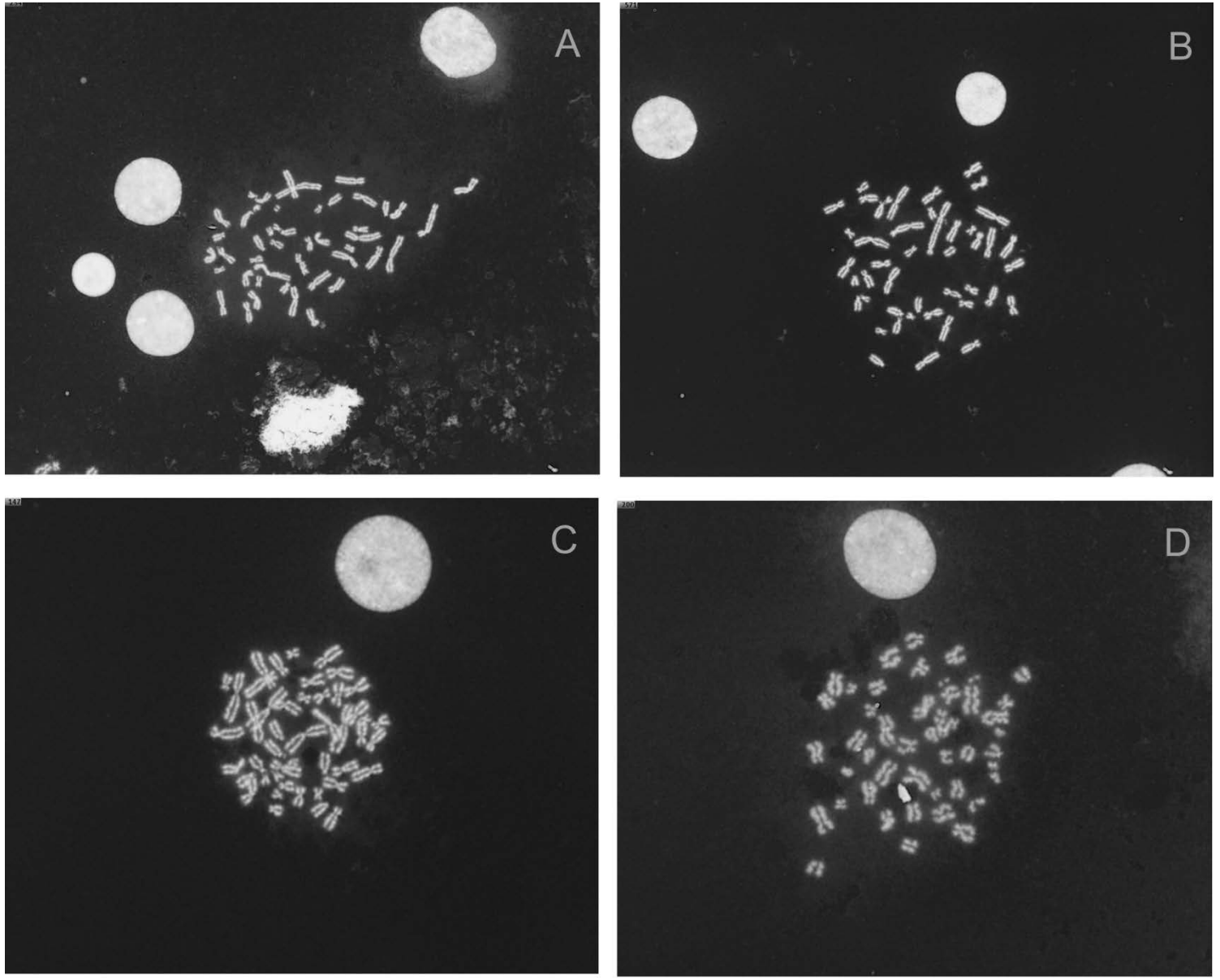
Segmented objects have colored contours. Red indicates DCs, yellow indicate chromosomes that were initially classified as DC, but eliminated based on FP morphology filters, green contours indicate MCs, and blue indicate objects failing ILL segmentation. No highlight: Nuclei, inseparable chromosome clusters, artifacts etc. Controls below image direct inclusion/exclusion of images from dose calculation.

Distribution of (Chromosome) Objects in a Set of Metaphases

- Metaphases must have objects in the user-specified range
  - Cell elimination of [<25-35, >60-75 objects] is recommended



ADCI can process images which have not been manually preselected. Thus, it was necessary to derive a set of filters to remove suboptimal images. Object count is one such filter.



Example:
Results of applying an image selection model

Examples of metaphase images in sample HCS05 (0.5Gy) ; unselected and selected by the 'group bin distance model , top ranked 250 images'*. **(A)** and **(B)** are selected images. **(C)** and **(D)** are images that have been eliminated by the model.

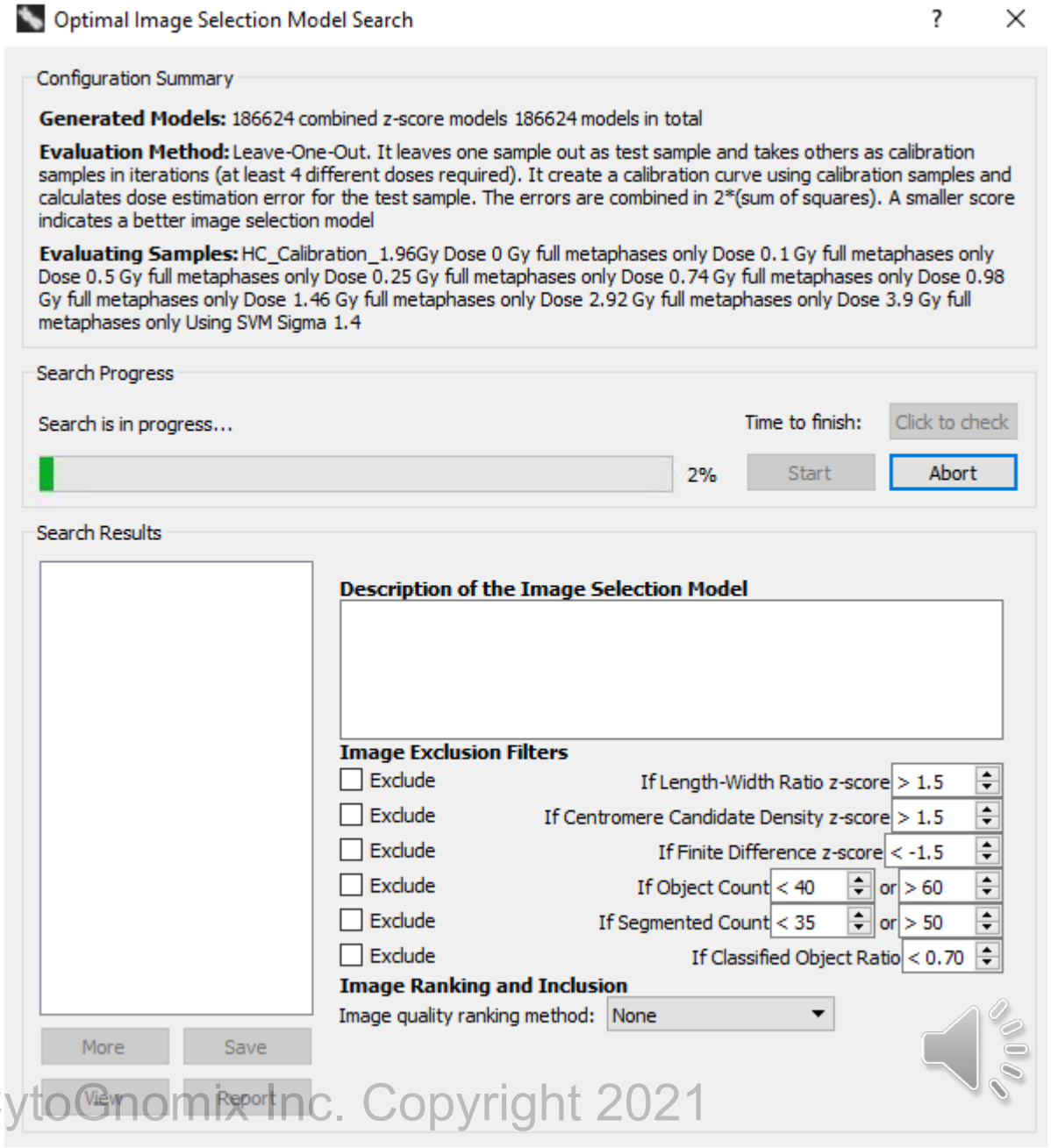
*Liu, Li *et al. F1000Res.* 2017, 6:1396

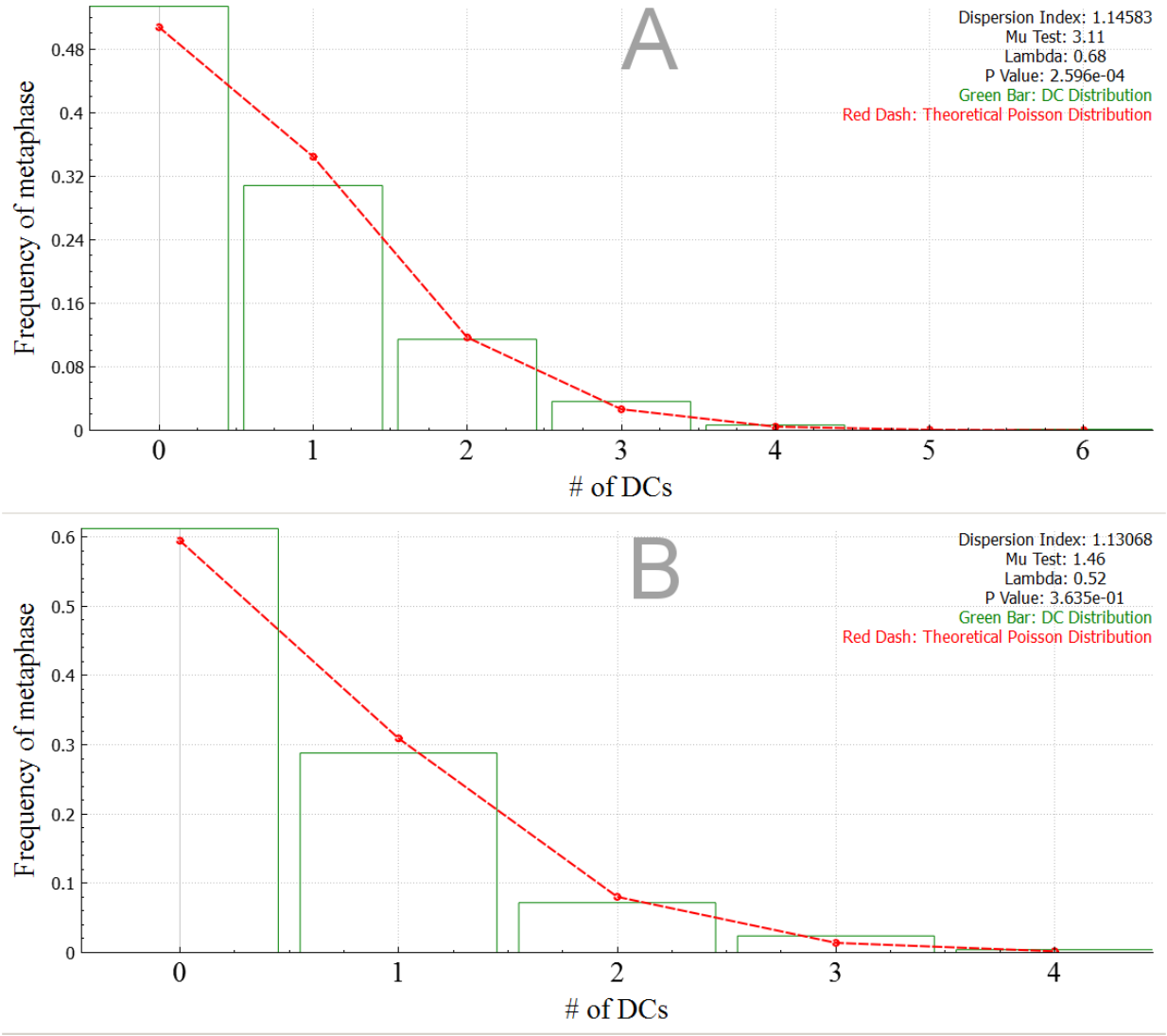
Demonstration: Determine an optimal image selection model

ADCI filters out suboptimal metaphase images by utilizing image selection models.

The effectiveness of a specific image selection model on calibration samples can be evaluated by observing the p-value of the Poisson fit for each sample, curve fit residuals, and leave-one-out dose estimation.

The "Optimal Image Selection Model Search" wizard automatically evaluates a large pool of image selection models and ranks them according to the selected evaluation method. This process takes significantly longer if the evaluation process is leave-one-out dose estimation as much of the evaluation process must be repeated with each calibration sample removed.



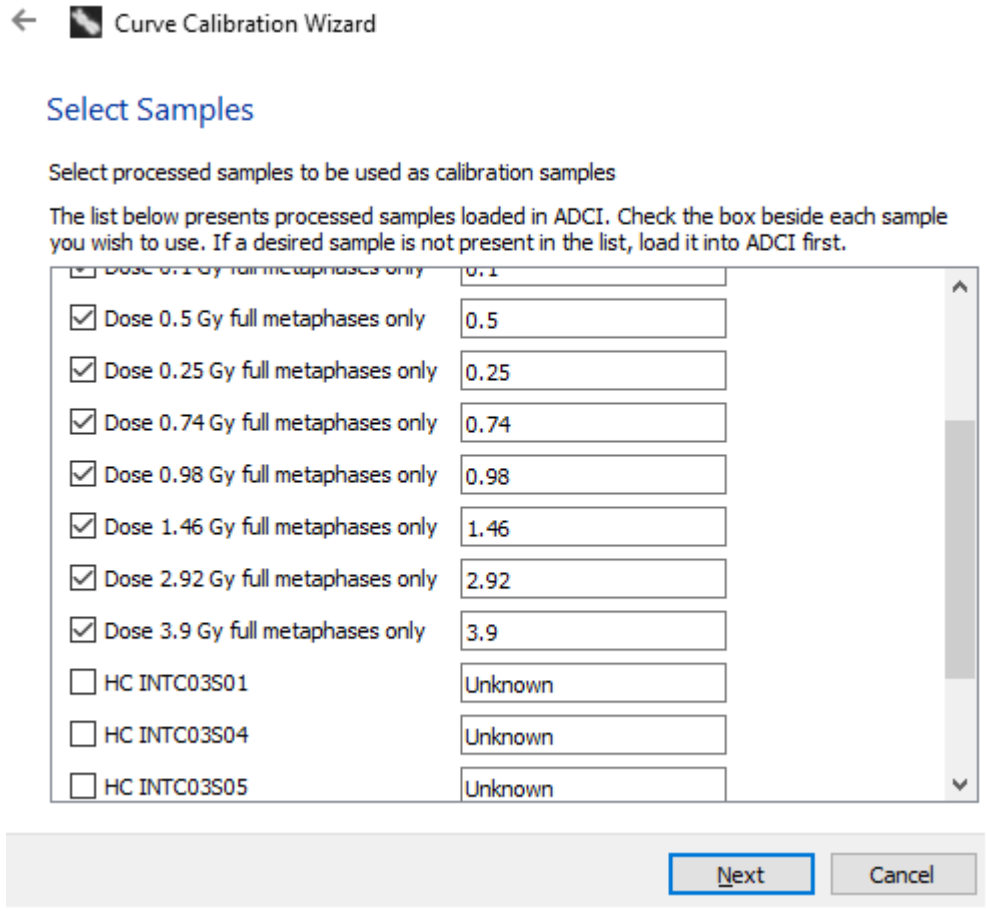


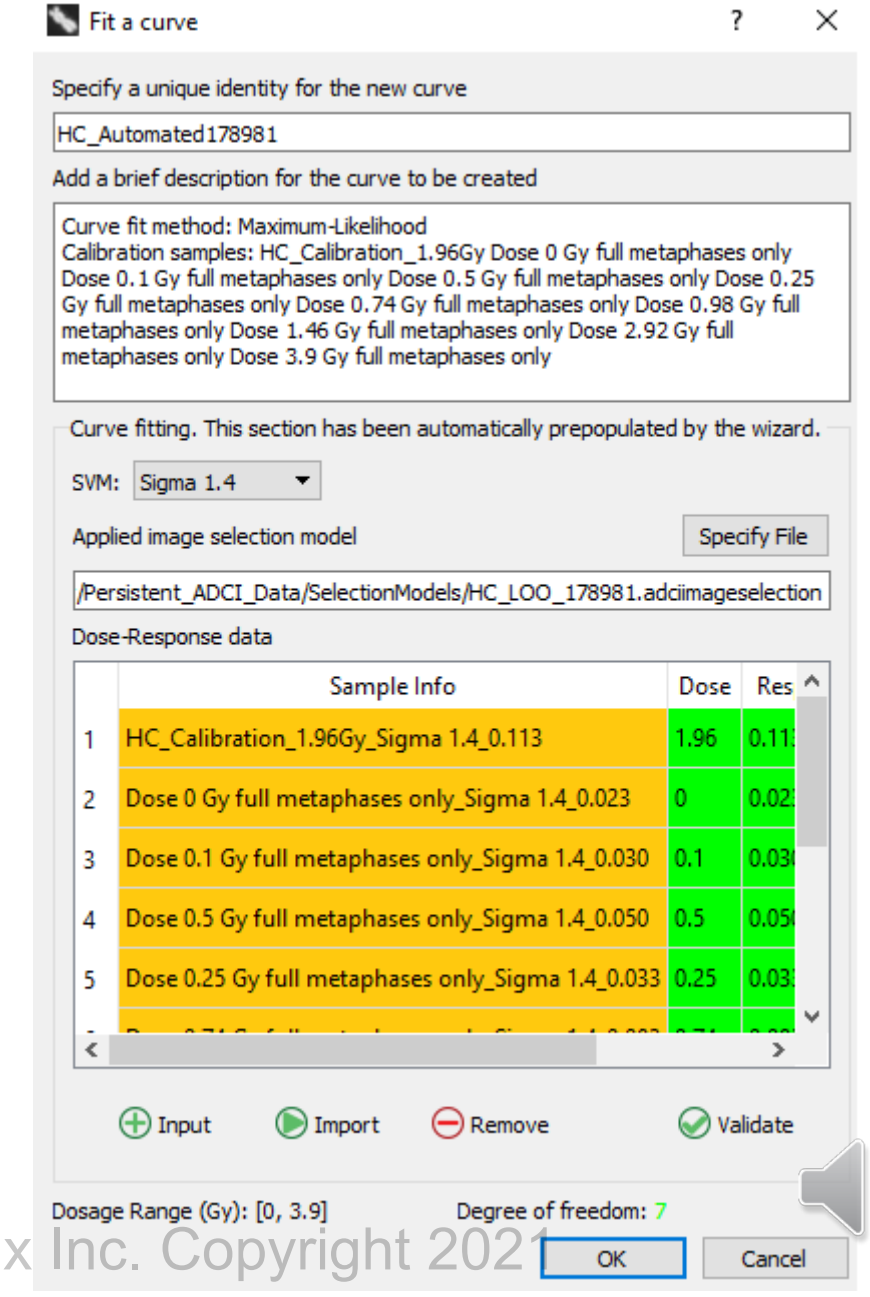
Poisson distribution of DCs in a sample prior to and post-image selection

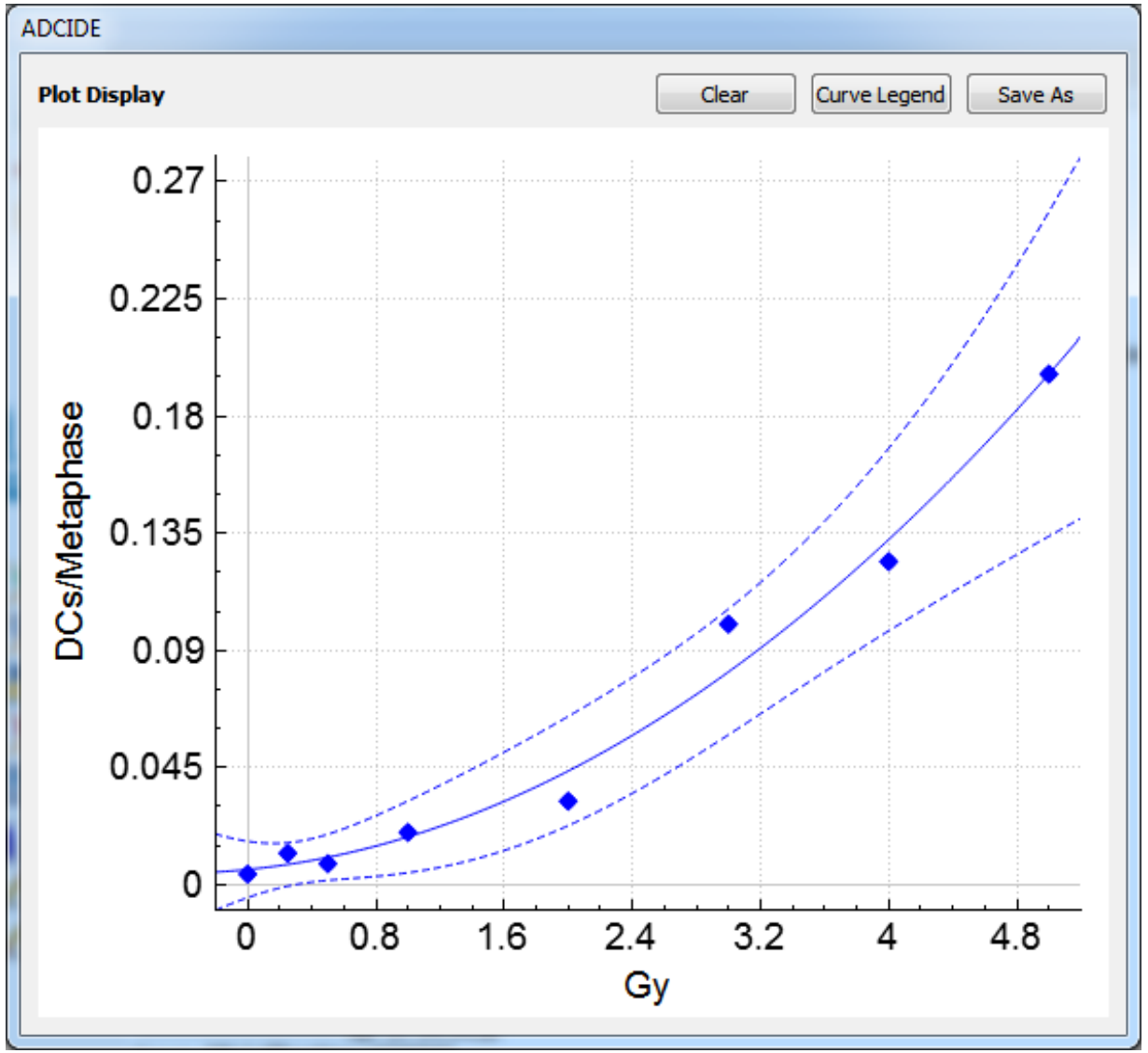
Screenshots of proportionate DC frequencies fit to Poisson distributions of Sample HC4Gy in ADCI. **(A)** All images are included (*no image selection*), **(B)** Only images selected by model (group bin distance, top 250 images) are included. The legend (top right) indicates the statistics of the fit to the Poisson distribution (Dispersion Index, Mu test, and Lambda) and the Chi-square goodness of fit test (p-value).

Demonstration: Create calibration curve

The calibration curve wizard is prepopulated with the physical dose of calibration samples if doses appear in the sample ID, otherwise the dose must be modified from “Unknown” to the known physical dose. Once created, calibration curves can be saved to the “Persistent_ADCI_Data” directory.







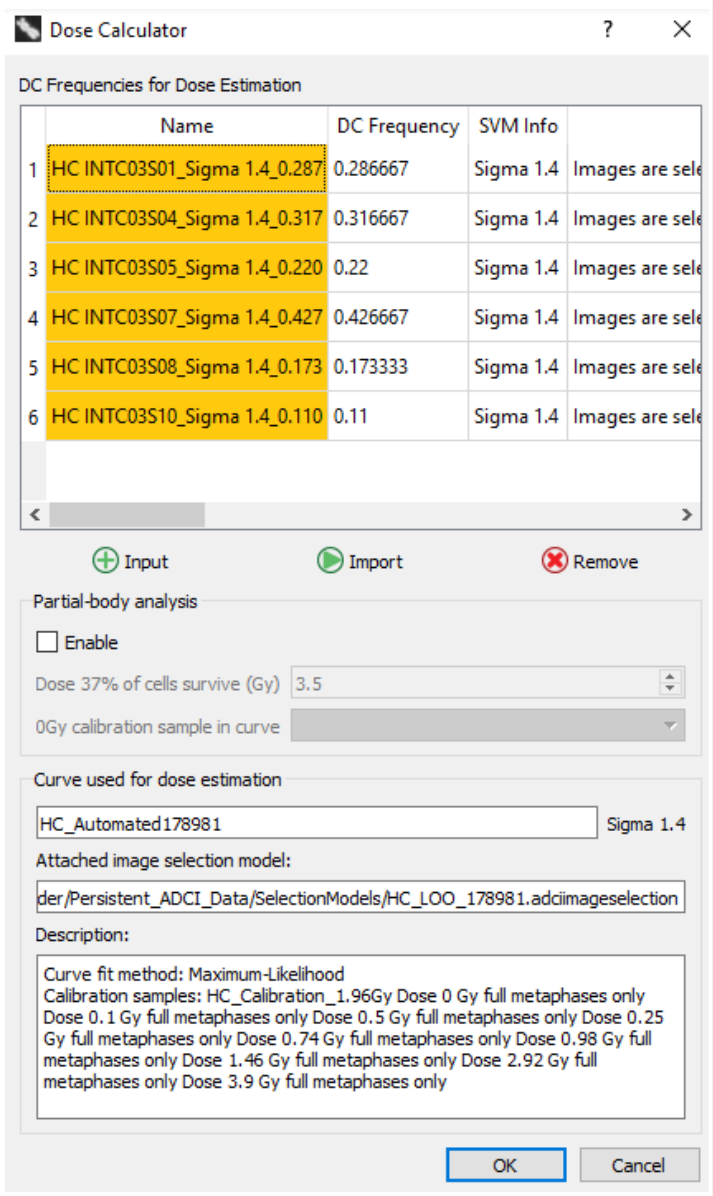
Screenshot of calibration curve (solid line) showing 95% confidence intervals (dotted lines). Parameters: Sigma = 1.5; Image selection model: top 250 ranked images sorted by combined Z-score method, for individual Z score tests I-VI, results weighted in proportion to: [5,2,4,3,4,1] for each test. This weighting combination is one of the optimal parameters found in a grid search on calibration data.

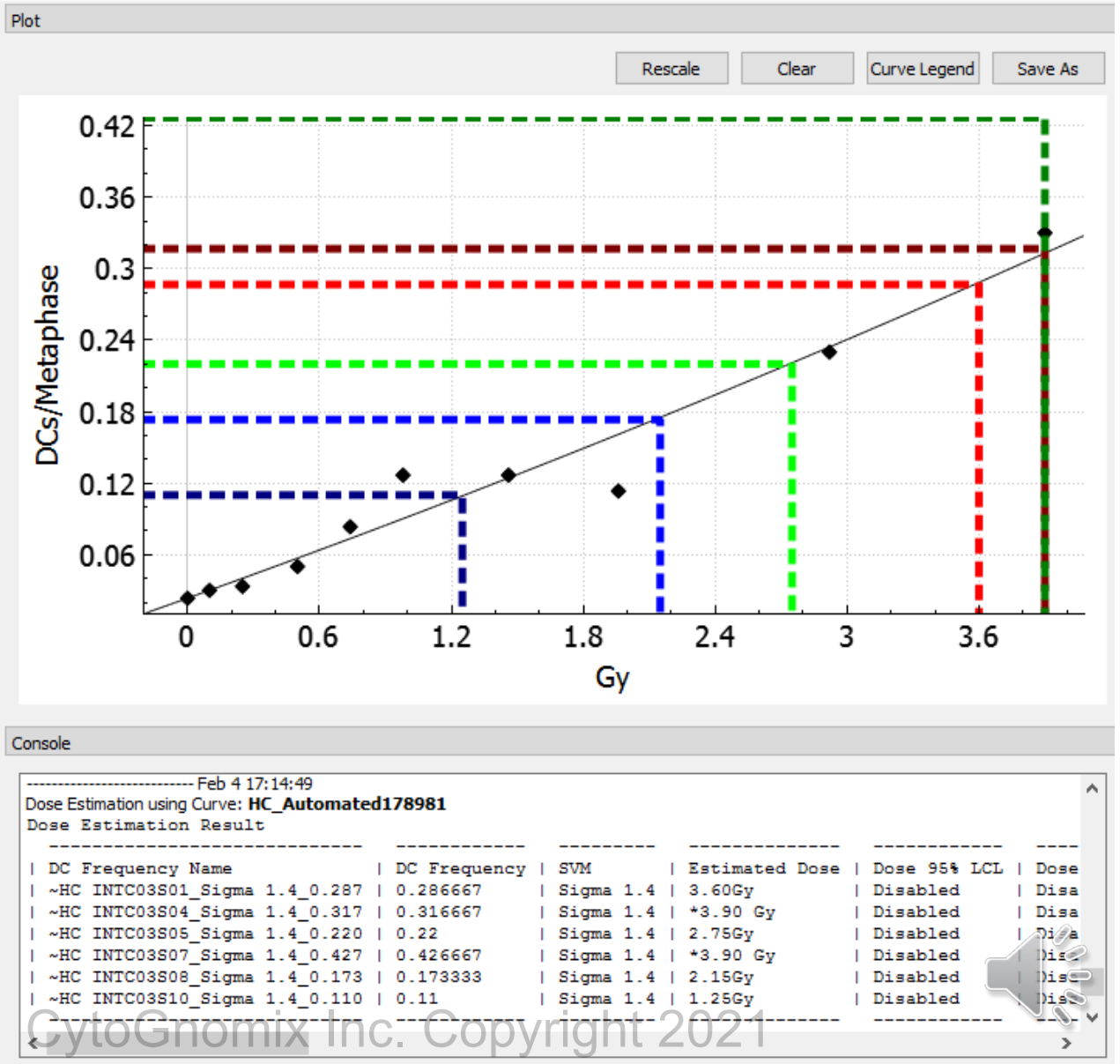
Demonstration: Estimate dose

The dose estimation wizard requires a set of samples of unknown dose, a calibration curve, and the associated image selection model.

The Plot and Console sections of the ADCI user interface are automatically populated when dose estimation is complete.

Results presented here mirror those described in *Rad. Prot. Dosimetry* **186(1)**: 42-47, 2019.



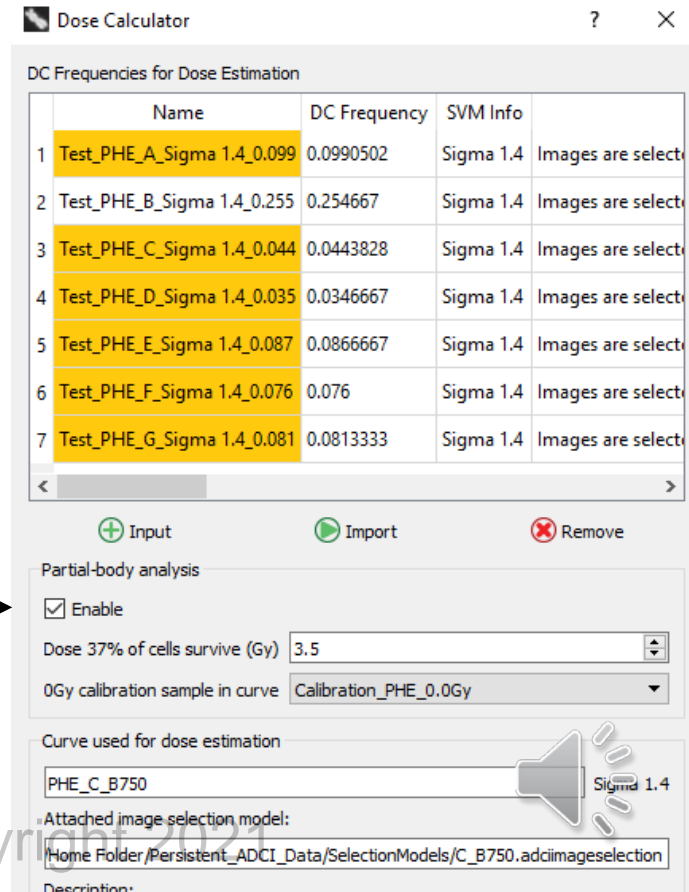


Demonstration: Estimating partial-body (PB) dose, confidence interval and fraction cells irradiated

DC Frequency Name	DC Frequency	SVM	Estimated Dose	PB Estimated Dose	PB Dose LCL	PB Dose UCL	PB Cells Exposed
~Test_PHE_A_Sigma 1.4_0.099	0.0990502	Sigma 1.4	1.85Gy	2.30Gy	1.20Gy	3.15Gy	80.71%
Test_PHE_B_Sigma 1.4_0.255	0.254667	Sigma 1.4	3.75Gy	4.20Gy	3.60Gy	4.75Gy	94.94%
~Test_PHE_C_Sigma 1.4_0.044	0.0443828	Sigma 1.4	0.50Gy	1.50Gy	*0.00Gy	2.85Gy	46.72%
~Test_PHE_D_Sigma 1.4_0.035	0.0346667	Sigma 1.4	*0.00 Gy	2.95Gy	*0.00Gy	4.85Gy	17.40%
~Test_PHE_E_Sigma 1.4_0.087	0.0866667	Sigma 1.4	1.60Gy	5.20Gy	4.00Gy	6.20Gy	47.68%
~Test_PHE_F_Sigma 1.4_0.076	0.076	Sigma 1.4	1.40Gy	2.25Gy	0.90Gy	3.20Gy	65.76%
~Test_PHE_G_Sigma 1.4_0.081	0.0813333	Sigma 1.4	1.50Gy	3.55Gy	2.30Gy	4.55Gy	52.41%

Often, radiation accident (and therapy) patients receive inhomogeneous exposures. In ADCI, partial-body dose estimation uses the 0Gy calibration sample in the selected calibration curve to reduce false positive dicentric chromosomes to estimate partial-body exposure and fraction of cells irradiated (*Int J Rad Biol.* **96(11)**: 1492-1503, 2020).

When this option is enabled, estimates of the partial-body dose and fraction of cells exposed are shown in the ADCI console and reports (PHE -E, -F, -G above), in addition to the standard output generated when analyzing homogeneously irradiated samples.



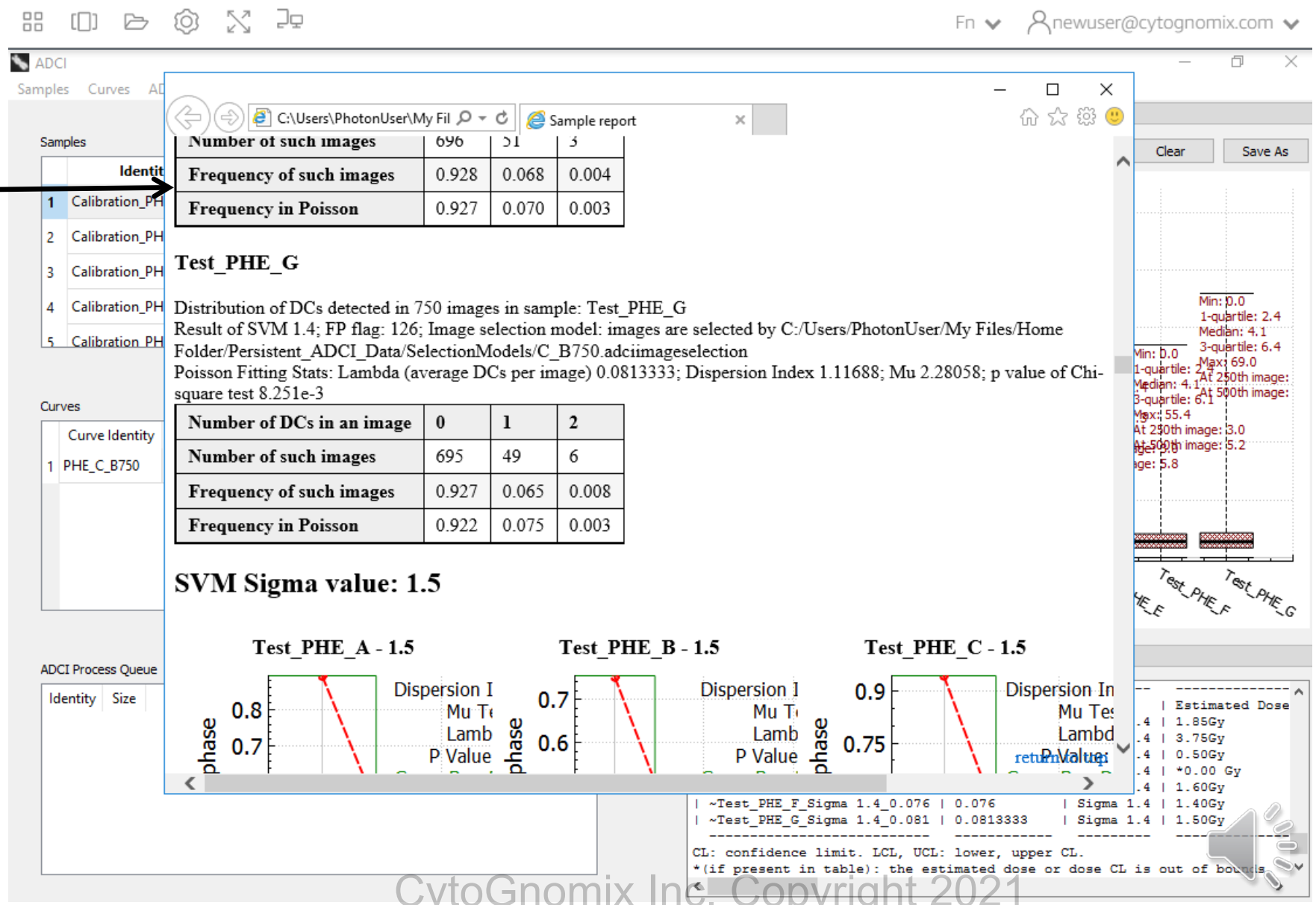
Demonstration: Generate and access written reports

ADCI_Online generates several report types including:

- Calibration curve
- **Sample**
- Optimal image selection model
- Dose estimation

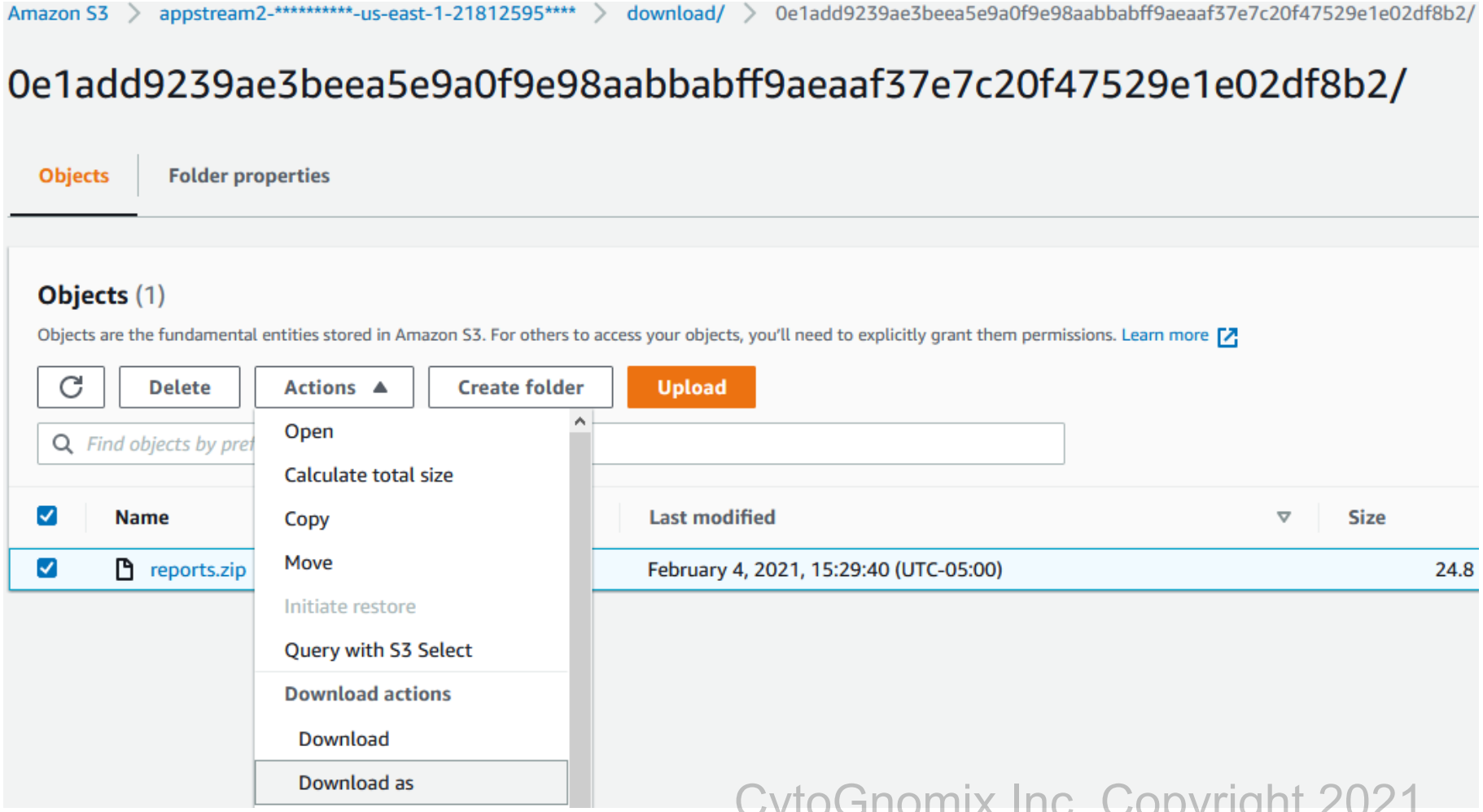
Reports are displayed in a web browser during the streaming session.

Alternatively, **saved reports can be downloaded from the user's AWS S3 folder at the end of the subscription period**.



Demonstration: Download reports

Upon completion of a subscription, any reports generated by the user are compressed as a single zip file and moved to a user-specific “download” directory where they can be downloaded to the Users local computer. The user accesses the data by accessing S3 as they did when uploading metaphase images.



Estimated time requirements

- Estimating sample processing time
  - When processing a set of samples, obtain an estimate of total processing time in minutes by summing image counts in all samples and dividing by 19.659 (the approximate number of images processed in one minute on ADCI_Online).
- Calibration sample processing
  - Assuming 7 calibration samples
  - 3 samples < 1 Gy (1500 images), 4 samples >= 1 Gy (750 images)
  - **7500** metaphase images in calibration samples
  - 7500 / 19.659 = 381.5 min (**6 hr, 22 min**)
- Automated Image selection model generation
  - An optimal image selection model must be determined only once for a set of calibration samples
  - Examine all model categories, curve fit residuals or p-value of Poisson fit evaluation modes: ~94 min (**1 hr, 34 min**)
  - Examine all model categories, Leave-one-out evaluation mode: ~233 min (3 hr, 53 min)
- Test sample processing
  - Assuming 7 homogeneously irradiated samples (700 images each) and 3 partially irradiated (1200 images each)
  - **8500** metaphase images in test samples
  - 8500 / 19.659 images = 432.4 min (**7 hr, 12 min**)
- Other categories of operations require little to no processing time, therefore the time required to perform them is limited by the operator's knowledge of the system. It is recommended to consult ADCI_Online documentation before a subscription begins. Time estimates below assume a moderately experienced operator:
  - Optional: Review of processed samples in the metaphase image viewer (**variable**)
  - Review of automatically generated image selection models, pre-existing, supplied models, or manual created models (**30 – variable min**). Manual creation and review of image selection models is an optional process and may require several hours.
  - Calibration curve generation (**5 – 30 min**)
  - Dose estimation (**5 – 30 min**)
  - Report generation and review (**10 – 120 min**)

Summary

- ADCI offers a cost-effective subscription-based service useful for radiation research, proficiency testing, inter-laboratory comparisons, and training.
- Security of data was considered at every step of the development process, with images and results encrypted and stored in an S3 Bucket directory accessible only to them.
- ADCI_Online can be rapidly scaled to meet “burst” requirements such as individuals in an emergency situation requiring processing of many samples.

ADCI weblinks

Introduction and access to demonstration version

https://radiation.cytognomix.com

Partnerships and contact e-mail address

info@cytognomix.com

How ADCI works (online manual)

https://adciwiki.cytognomix.com

Dicentric chromosome classification by machine learning

https://cytobiodose.cytognomix.com

ADCI protocol in the Journal of Visualized Experiments (JoVE)

https://doi.org/10.3791/56245

To obtain ADCI

https://radiation.cytognomix.com/quoterequest.php