GALEAS Software Instructions for Use

Table of contents

GALEAS™ nonacus

Introduction

The following guide explains how to get set-up on the Nonacus GALEAS software and how to submit samples for processing. There are some steps you will need to do just once and then others each time you want to submit a batch of samples for processing.

One Time Steps	Per batch of samples
• Create a MyNonacus account	• Login via Desktop sample upload application
• Download the Desktop sample upload application	• Create an ARF for the batch of samples
• Download example ARF template	• Use the Dekstop application to upload samples
	• Read the Batch notification email
	• Login in to MyNonacus
	• Download the results of the sample processing

Create a "My Nonacus" Account

If you haven't already done so, please open your web browser (e.g. Edge, Chrome, Safari) go to **www.nonacus.com** and register for "My Nonacus" by clicking the My Nonacus logo of the top right of the screen

⎙ My Nonacus

After you have registered, you can request an access to GALEAS in your Account by sending a mail to **support@nonacus.com.** Your account will be associated with the GALEAS software, which enables you to upload patient samples (in FASTQ format) and then download the results.

Login to MyNonacus

Open your web browser (e.g. Edge, Chrome, Safari) and goto **mynonacus.nonacus.com,** this gets you into the My Nonacus portal, where you can design custom Panels, view Catalogue panels and analyse patient samples.
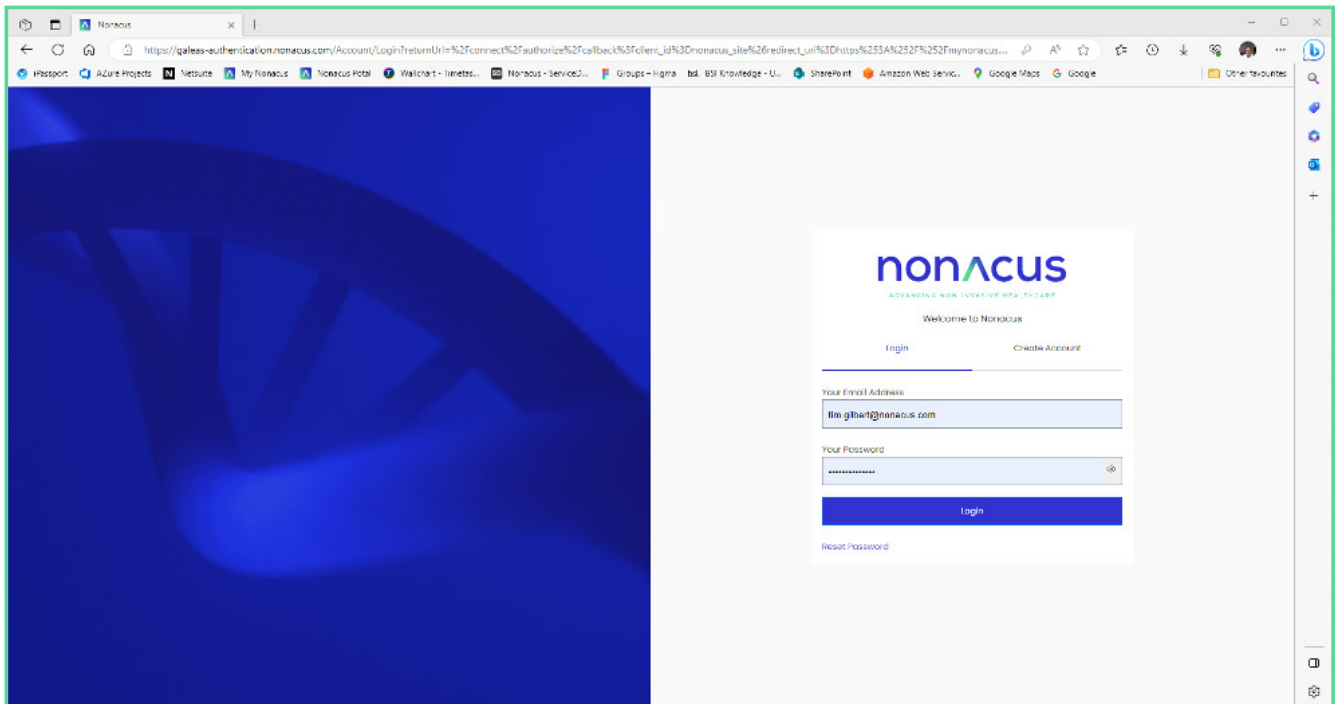
Download the Desktop sample upload tool from the resources tab

In order to upload FASTQ files, you will need the Desktop file uploader which is available for download to your computer from the Resources tab. Click on the link to download and install the File uploader application onto your PC.
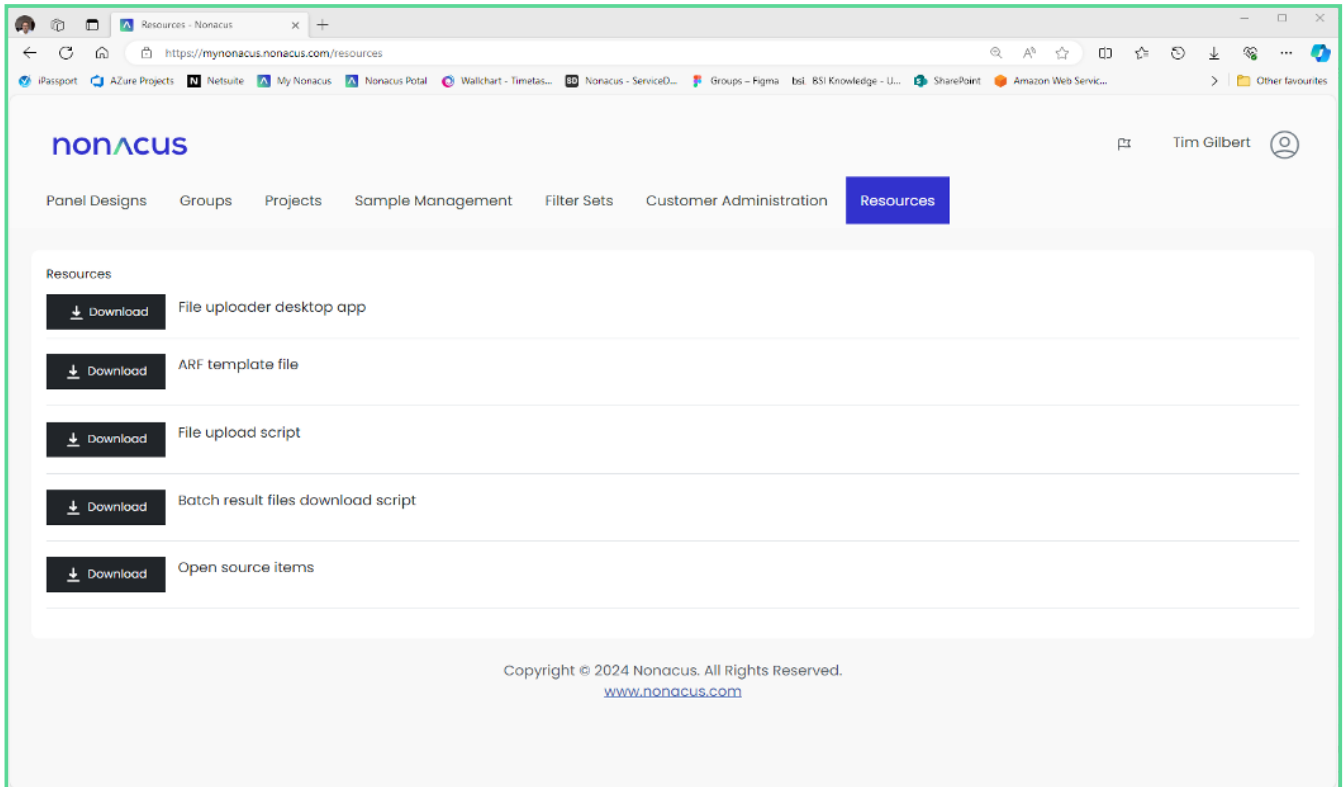
Download example Analysis Request File (ARF) file

In a similar way that a sample sheet tells a sequencer how to process the physical samples, the ARF file lets the GALEAS software know which FASTQ's to upload and how to associate them with a patient in the system. An example of a ARF file can be downloaded from the system by clicking on the resources tab and then clicking on "ARF Template file" 'download' link.
The format of the ARF is as follows;

|	
SampleID*	A mandatory field, that must be unique, to identify the sample in the system. The FASTQ filename must contain the sampleID in it, to enable the application to 'find' the FASTQ file(s) and associated it with this sample when the filenames are not explicitly provided in the 3 filename columns below (FASTQ Forward filename etc.). The system will automatically work out whether the r1, r2 (and r3) are forward, reverse or UMI files.
Sample Type*	A mandatory field which must be one of the following; • FF • FFPE • Blood • Urine Pellet • Saliva Depending on the Panel being used
PanelID*	A mandatory field, this is the panel associated with the sample, the ID's are as follows; 1821 for GALEAS Bladder 1969 for GALEAS Hereditary Plus 1911 for GALEAS Tumor
PatientID*	A mandatory field this is the unique ID for the patient – it can be an existing ID, in which case the sample will be associated with that patient in the system, or it can be a new ID, in which case a new patient will be created with the details supplied in the ARF.
Lane	Optional field to be used when the sequencer produces FASTQ files for each lane, in which case the filenames below must be provided for each lane specified here (lanes 1-4). If the sequencer produces multi-lane files (most common) then this field can be left blank.
FASTQ Forward filename (.1)	Optional field which specifies the filename of the FASTQ file for forward reads e.g. S0086.1.fastq.gz
FASTQ Reverse Filename (.2 or .3)	Optional field which specifies the filename of the FASTQ file for reverse reads e.g. S0086.3.fastq.gz
FASTQ UMI filename (.2)	Optional field which specifies the filename of the FASTQ file for UMI e.g. S0086.2.fastq.gz
SampleDate	Optional field to specify when the sample was taken
Sample Description	Optional description field for the sample
DNA ng/ul	Optional field specifying the amount of DNA in ng per ul
Elution vol	Optional field specifying the elution volume

...cont:

Plasma	Optional field specifying the plasma volume
Project*	This field is optional if a patientID is already associated with a Project. If the patientID is new (i.e. not an existing patient) then this is a Mandatory field specifying the project to associate the patient with. The project must be created in the system prior to sample upload & processing.
First Name	Optional field specifying patient first name
Surname	Optional field specifying patient surname
Date of Birth	Optional field specifying the DoB of the patient
Cancer Diagnosis/Conditions	Optional field, specifying the disease/condition of the patients
Comment	Optional field for a comment
Clinician Forename	Optional field for the requesting clinician first name
Clinician Surname	Optional field for the requesting clinician surname
Clinician Address Line 1	Optional field for the requesting clinician address (1st line)
Clinician Address Line 2	Optional field for the requesting clinician address (2nd line)
Clinician City	Optional field for the requesting clinician city
Clinician Country Name	Optional field for the requesting clinician country
Clinician Postcode	Optional field for the requesting clinician postcode
Received Date	Optional field specifying the date the sample was received at the lab

Create a project.

Projects are a way of grouping patients together in the GALEAS system. Projects are associated with a customer, so any user belonging to a customer can see all the projects of that customer. A Patient can only be in one Project.

There is a 'default' project for a customer, but we recommend utilising the project concept to keep the patients organised.
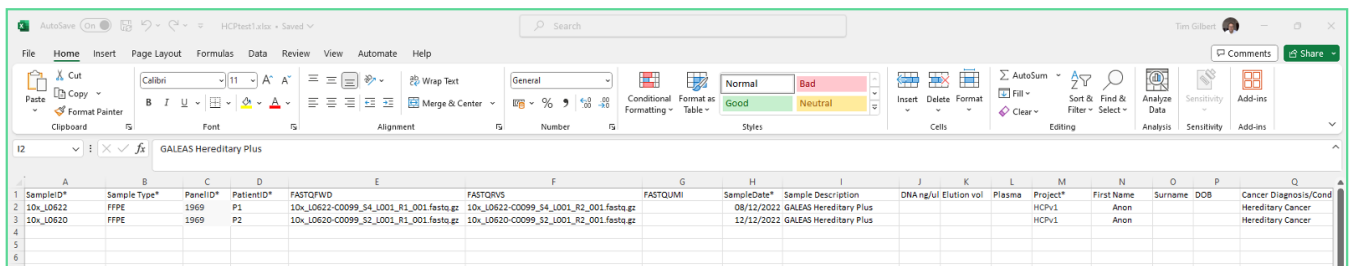
To create a project, click on the "Projects" tab, you will see a list of any existing Projects, and a list of Patients for the project that was last selected, if applicable. Click on the "+ Add Project" button to create a new project.
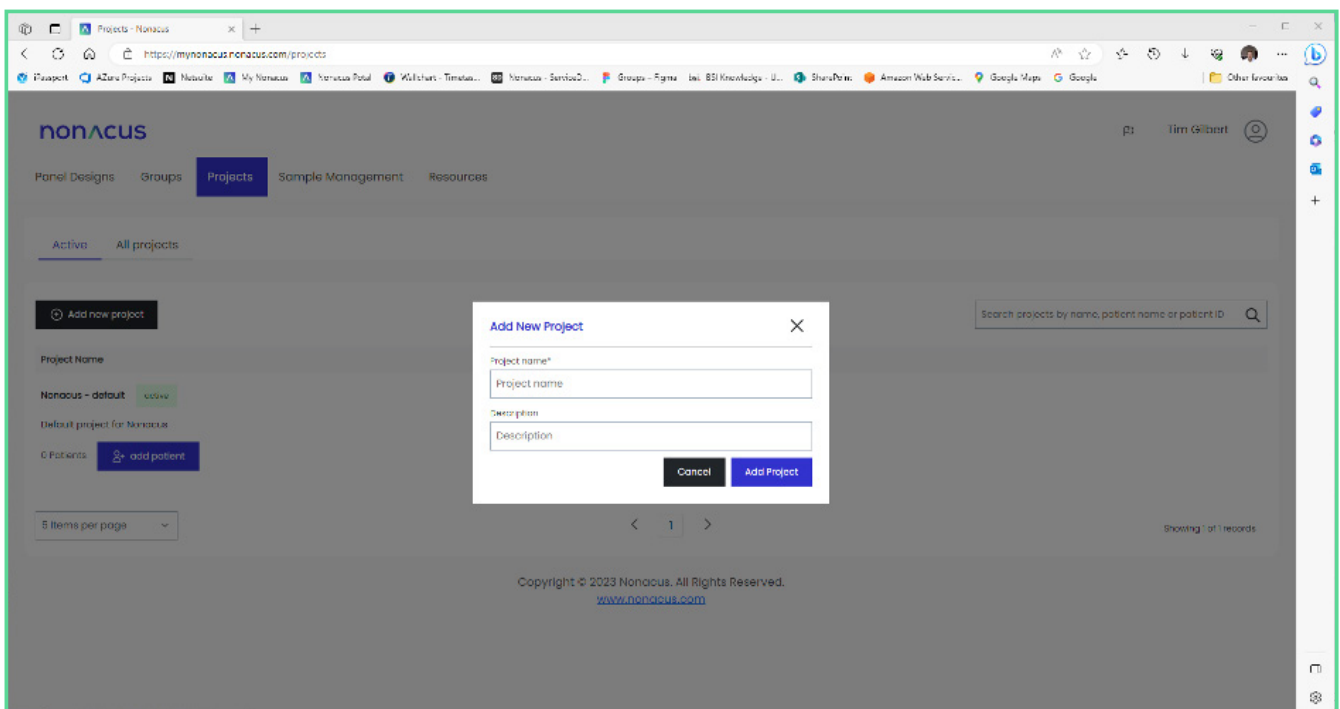
Create an ARF for your samples

The ARF file lets the GALEAS software know which FASTQ's to upload and how to associate them with a patient in the system. See section "Download example Analysis Request File (ARF) file" for an explanation of the ARF format.

Important note: No personally identifying information (PII) for the patient should be provided in the ARF file. Only non-identifiable ID's should be used.
An example ARF is shown below.

Using the template as a guide, complete the ARF for the batch of samples you want to process. The ARF will be used by the Desktop upload, which can be downloaded following instructions provided in section "Download the Desktop sample upload tool from the resources tab"

Login via the Desktop Sample upload tool

Using the desktop sample upload application requires you to login. Simply enter your username (your email address) and your password and then press login.

The desktop application will then allow you to upload a single sample ("Single Sample Upload tab") or a batch of samples using the Batch Upload tab – see next section "Choose the 'batch-upload' tab for multiple files using the ARF".
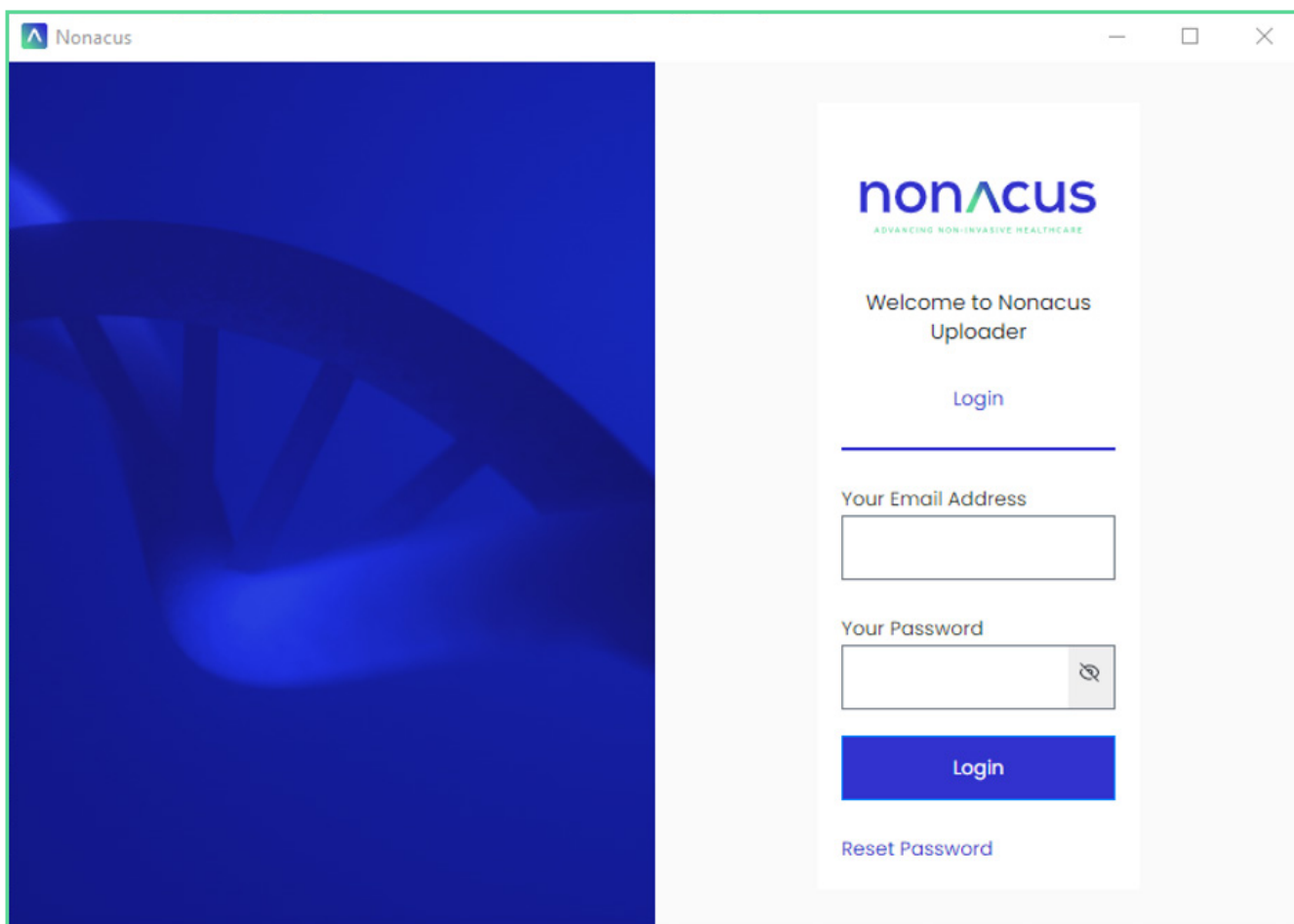
# Uploading a Batch of Samples using Desktop Application

Choose the 'batch-upload' tab in the Sample upload application.
Firstly select your completed ARF by clicking "Choose ARF" button and using the file-finder to navigate and select your ARF. Then click the "Choose files" button to select the location of the sample FASTQ files.

Finally click on "Upload" button. The GALEAS application will then perform some checks, such as ensuring all the samples you've listed in the ARF have corresponding FASTQ file(s) in the location you've specified. If there are any errors, a message will be displayed informing you what the issue is and how to fix it.

Uploading even a small number of FASTQ files can take hours, depending on the nature of the samples (e.g. small panel versus Exome) and your broadband speed. Please leave the application running until it says it has completed the upload.

Once the samples have been uploaded, the GALEAS system will process the files, and again this process could take several hours. You will receive an email notification when the processing is complete – see next section "Wait for the batch notification email".
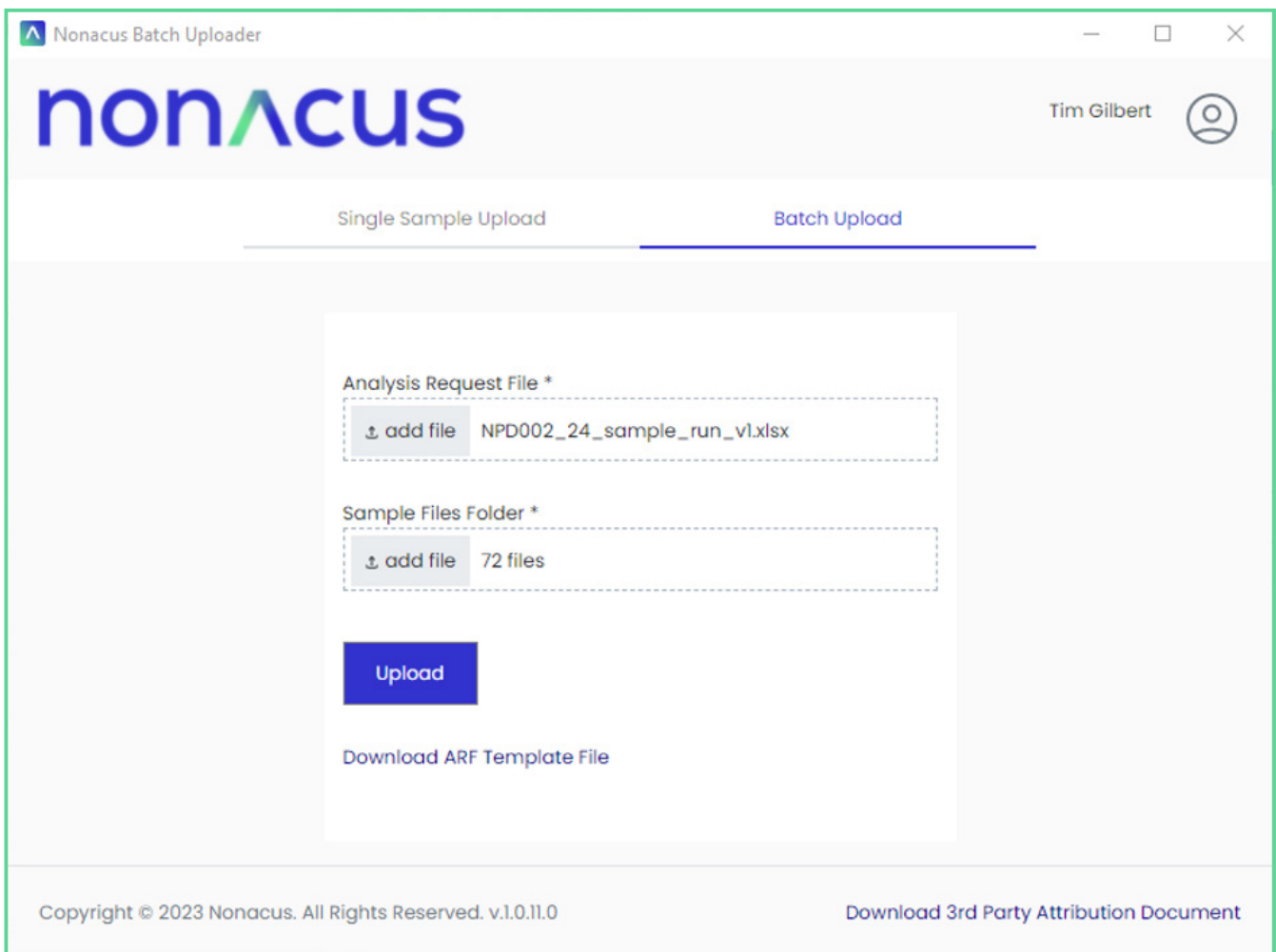
Wait for the batch notification email

Once the GALEAS solution has finished processing the batch of samples, it will send an email to you. In the email is a link that will take you direct to the "Batch download pages" in GALEAS (you may need to log in, but the link will take you to the correct place after successfully logging in), where you can download the results i.e. the report.
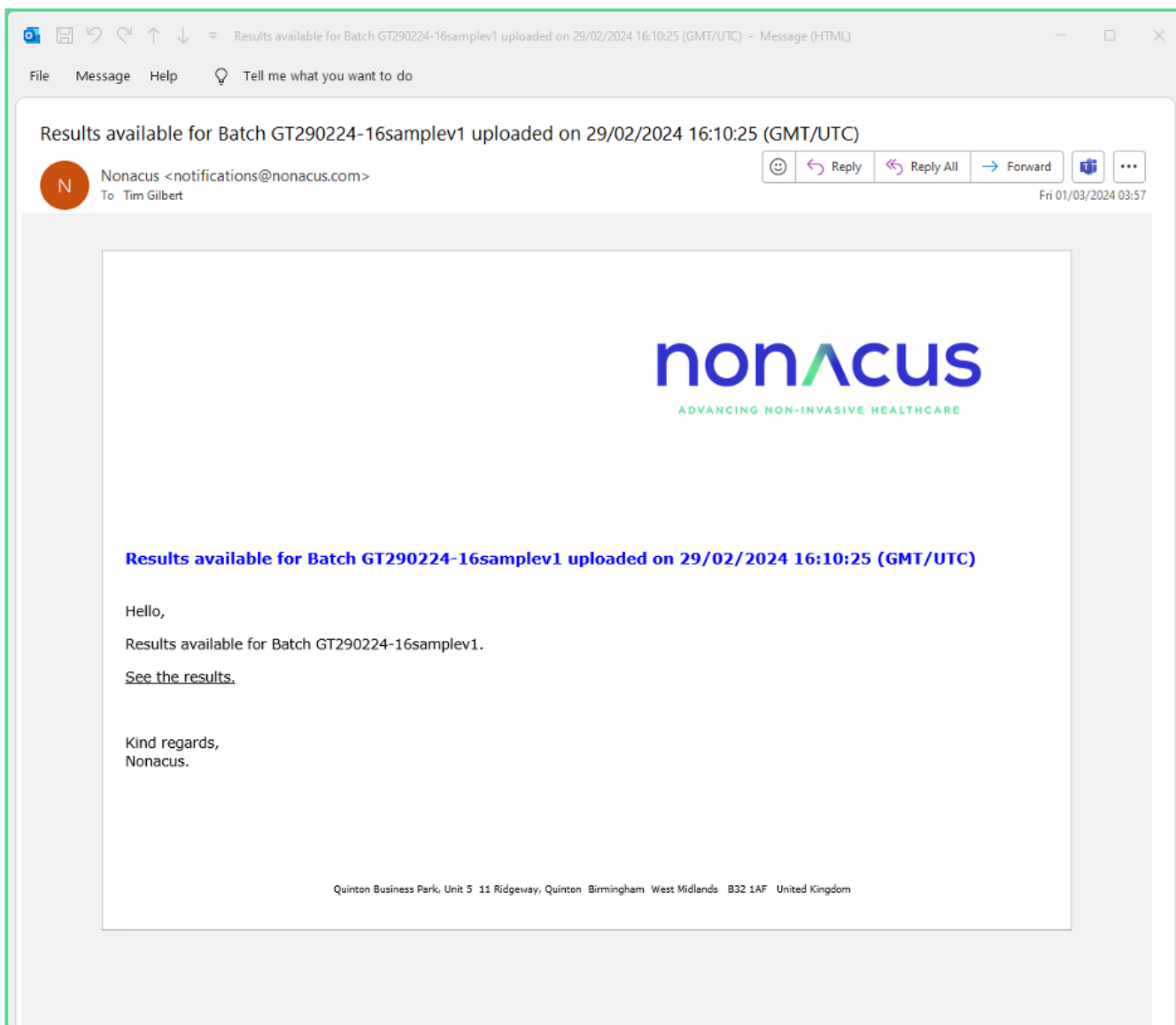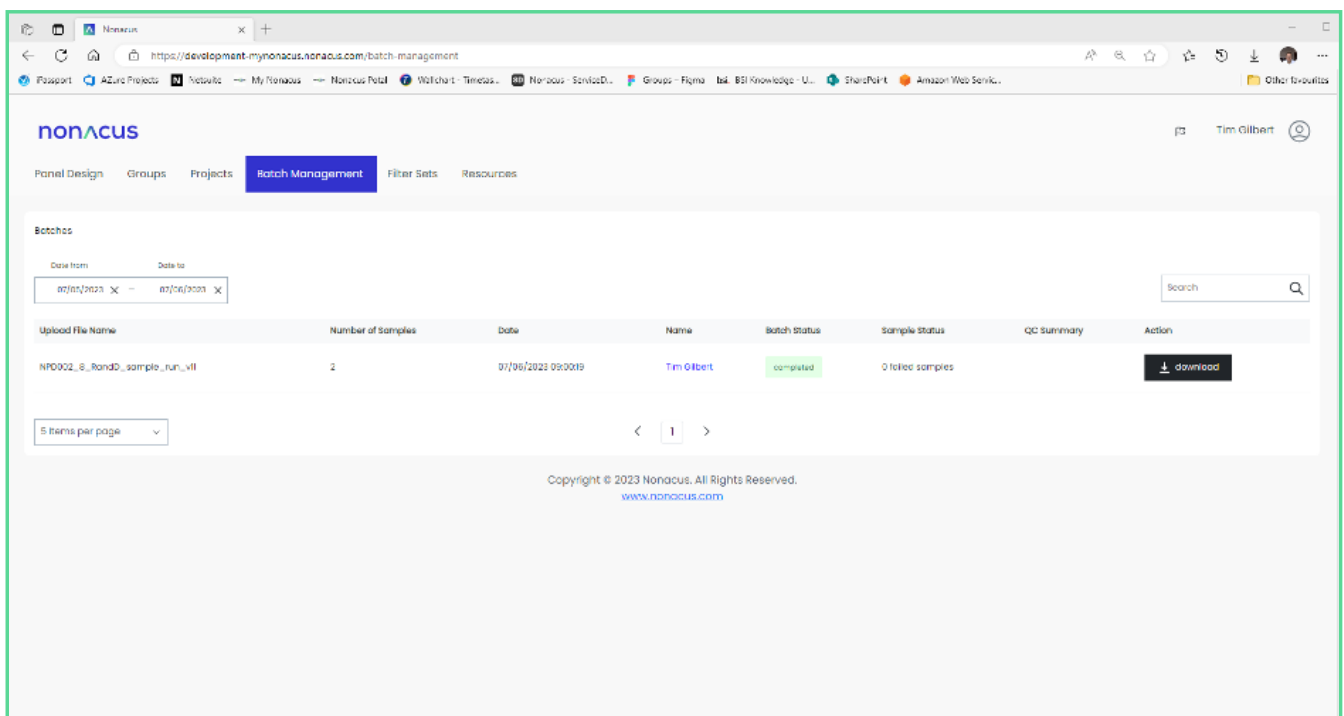
Download the result for the batch(es)

The link in the email will have taken you to the "Batch Management" tab in the GALEAS software. This page shows you all the batches that have been processed for your customer, with each row being a sample batch processing run.

The 'Download' button will download all the results for each of the samples in the batch, to your usual 'downloads' folder on your PC.

For each successfully processed sample, the download contents will depend on the panel being used – see the annex for a description of the contents for GALEAS Bladder, Hereditiary Plus and Tumor.

Uploading a Batch of Samples using Python script

To allow for the automation of sample processing, a Python script is available that can be called from the command line with a path to an ARF file and a path to a folder containing the sample files.
Note: The script is designed to work on both Windows and Linux machines

Pre-requisites
Install Python (v3 or later) by going to **https://www.python.org/downloads/** and following the installation instructions for your operating system.

Install dependencies

1. **python -m pip install requests**

2. **pip install openpyxl**

Get the Script
The script is customer specific for security reasons, so to get your copy you need to

1. Login to the customer site

2. Go to the resources tab

3. Click the download button

Run the Script
Once you have prepared the ARF file and sample files you can run the script as follows;

python ARFUpload.py -a "path_to_ARF_file.xlsx" -s "path_to_sample_files_folder"

Sample Results Download Script

The download script allows users to download the results of a specific batch run, assisting them in automating their NGS workflows.

Setting up

1. Install Python by going to **https://www.python.org/downloads/** and following the installation instructions for your operating system.

2. Install dependencies: python -m pip install requests

3. Log in to the MyNonacus site

4. Download the script from the 'Resources' tab:

This will save the file 'GaleasDownload.py' into your Downloads folder.

Running the script

- Move the python script into the desired folder. The script will save files into the same folder as the script is saved in - e.g.: if the 'GaleasDownload.py' file is moved into the Documents folder, the downloads will also be saved into the Documents folder.
- The batch that you request to download must be in a completed state. Batch files in a processing state can not be downloaded.

Downloading sample results

Use the following command to run the script:

python GaleasDownload.py "<batch name>" <file types>

<batch name>: the name of the ARF that was uploaded for this batch.
<file types>: a list of file types to be included in the download (case insensitive).
The following file types are supported:
- PDF
- VCF
- BAM
- JSON
- QC
- MAF
- TSV
- TXT
- ALL (include all files)
Example command:

python GaleasDownload.py "Batch upload 1" PDF JSON MAF

Description of Download contents

GALEAS Bladder
The download bundle for a GALEAS Bladder sample contains the following;
- PDF of the report
- JSON encoding of the report
- multiQC.html – web page containing a large variety of sample QC data obtained from multiQC

GALEAS Hereditary Plus
The download bundle for a GALEAS Hereditary Plus sample contains the following;
- BAM and BAM index file (ends with .bam and .bam.bai)
- SNV/Indel germline VCF and VCF index (ends with .germline.vcf.gz and germline.vcf.gz.tbi)
- Ensembl VEP annotated SNV/Indel germline VCF (ends with .germline.vep.vcf.gz)
- Germline CNV VCF and VCF index (ends with germlincnv.vcf.gz and germlinecnv.vcf.gz.tbi)
- Log2 Ratio CNV plots (ends with germlinecnvplots.zip – one .html file per gene with CNV calls in the zip)
- multiQC.html – web page containing a large variety of sample QC data obtained from multiQC

Note: SNV/Indel calling is by Sentieon, with extra annotations by Ensemble VEP, CNV calling is by cn.mops, ExomeDepth and Sentieon CNV callers.

GALEAS Tumor
The download bundle for a GALEAS Tumor sample contains the following;
- Log2 ratio CNV scatter image (ends with Log2.scatter.pdf)
- BAM and BAM index file (ends with .bam and .bam.bai)
- SNV/Indel germline VCF and VCF index (ends with .germline.vcf.gz and germline.vcf.gz.tbi)
- MSI scores (ends with .msi)
  ◦ 3 data points - Total Number of Sites, Number of Somatic Sites, MSI percentage
- SNV/Indel somatic MAF (mutation annotation format) file (ends with .somatic.maf)
- SNV/Indel somatic VCF and VCF index (ends with .somatic.vcf.gz and somatic.vcf.gz.tbi)
- Ensembl VEP annotated SNV/Indel somatic VCF (ends with .somatic.vep.vcf.gz)
- Somatic CNV VCF and VCF index (ends with somaticCNV.vcf.gz and somaticCNV.vcf.gz.tbi)
- Somatic SV VCF and VCF index (ends with somaticSV.vcf.gz and somaticSV.vcf.gz.tbi)
- TMB score (ends with TMB.tsv)
  ◦ 4 data points - Variants remaining after filtering, Total length of target regions (bp), Tumour mutational burden (Variants/Mb), Panel size warning (true/false)
- multiQC_report.html – web page containing a large variety of sample QC data obtained from multiQC