Using the 10x Genomics pipeline **Cell Ranger ARC Count v2.0.2** (<https://support.10xgenomics.com/single-cell-multiome-atac-gex/software/pipelines/2.0/using/count>)

1. **Download sequence data from Novogene to Alabama Supercomputer (ASC)**
   1. On ASC:

wget <download link>

1. **Upload sequence data to 10x Genomics Cloud server**
   1. <https://support.10xgenomics.com/cloud-analysis/uploading-fastqs>
   2. On ASC, first download the 10x Genomics Cloud CLI:

curl -f -o txg-linux-v1.3.1.tar.gz \

https://cf.10xgenomics.com/cloud-cli/v1.3.1/txg-linux-v1.3.1.tar.gz

* 1. Then unpack it:

tar -zxvf txg-linux-v1.3.1.tar.gz

* 1. On ASC: Navigate to “txg” folder (from 10x Genomics) and enter:

txg.exe fastqs upload --project-id <project ID> <path to FASTQs>

1. **Create a custom reference file for 10x Genomics Cloud analysis**
   1. NOTE: the pipeline used in the 10x Cloud Analysis (Cell Ranger ARC Count v2.0.2) requires a custom reference sequence created from ‘cellranger-arc mkref’ (<https://support.10xgenomics.com/single-cell-multiome-atac-gex/software/pipelines/latest/tutorial/mkref>)
      1. On ASC, create a .config file like the following: (“Ref.config”). This will be used for the following step using cellranger-arc mkref.

{

organism: “Canis\_lupus\_familiaris”

genome: [“Dog10K\_Boxer\_Tasha”]

input\_fasta: [“Ref.fna”]

input\_gta: [“Ref.filtered.gtf”]

}

* + 1. On ASC, run the following script (large, 12 cores (8%), 36gb memory (34%), takes ~1hr):

#!/bin/sh

source /opt/asn/etc/asn-bash-profiles-special/modules.sh

module load cellranger-arc/2.0.2

cellranger-arc mkgtf \

Ref.gtf \

Ref.filtered.gtf \

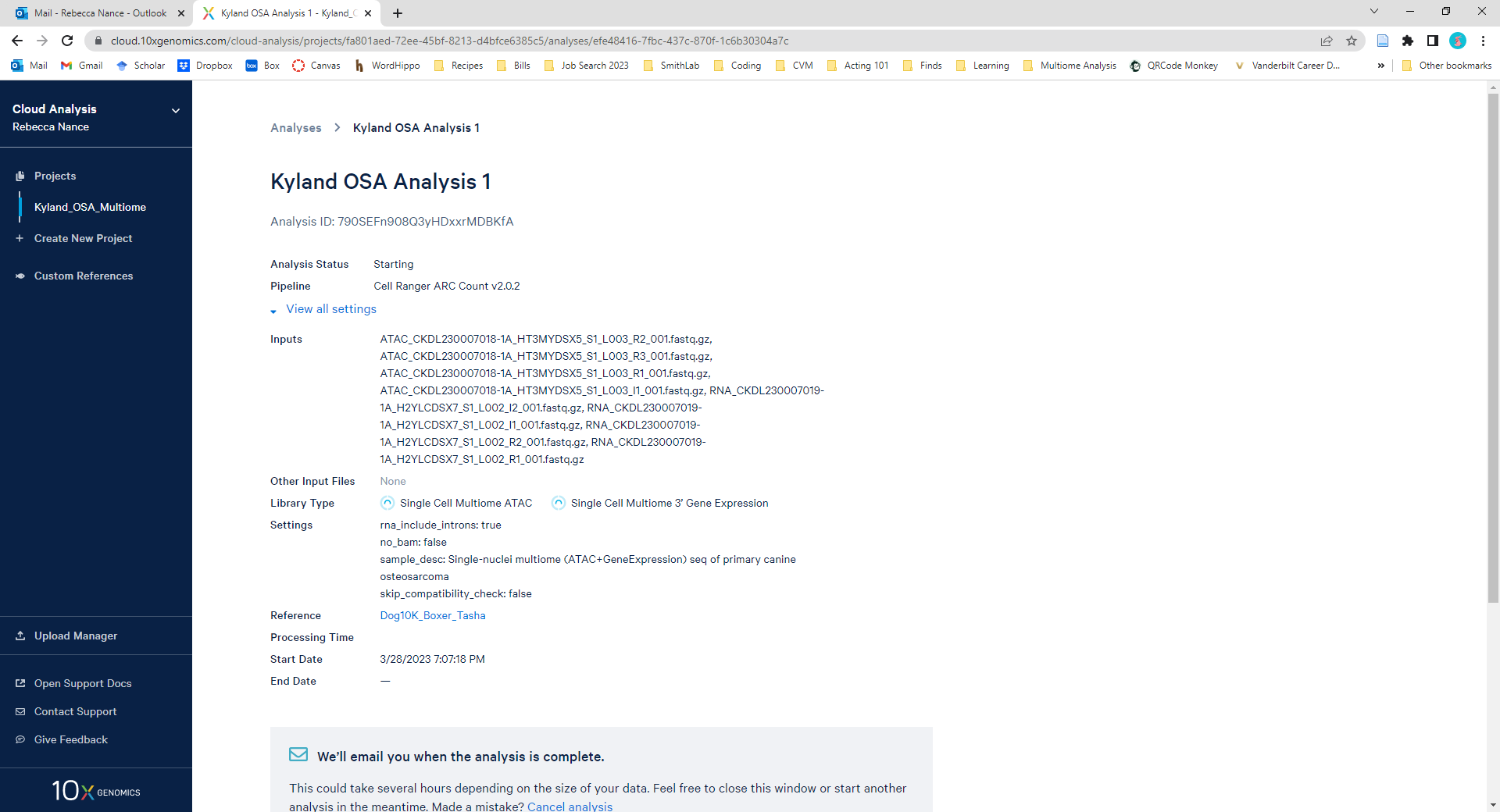
--attribute=gene\_biotype:protein\_coding

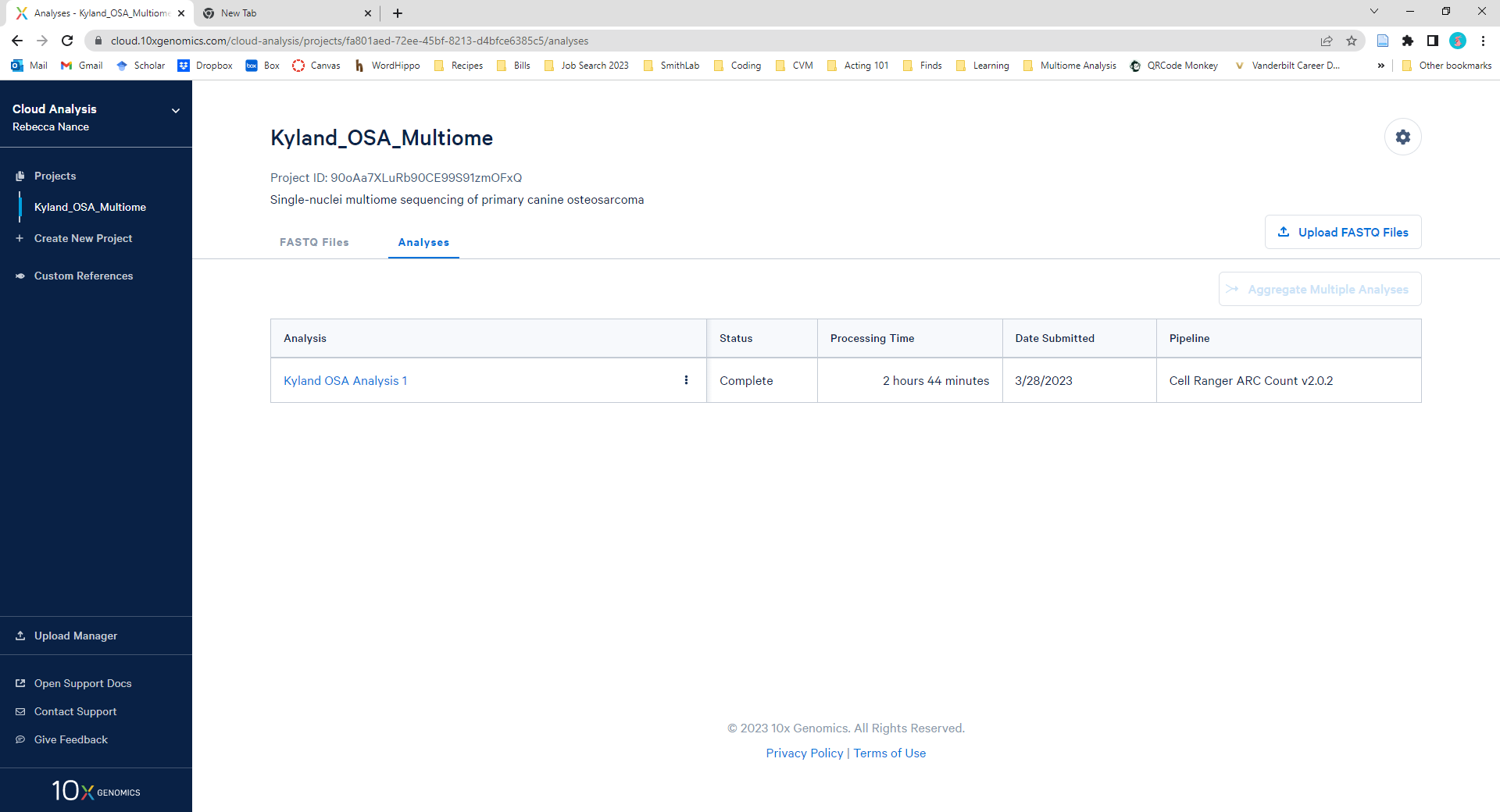
cellranger-arc mkref --config=Ref.config

* 1. Upload the reference directory you just made to 10x Cloud
     1. Navigate to the 10x Cloud Upload Folder (txg-linux-v1.3.0) and enter:

./txg references upload <path to reference>

1. **Start new analysis on 10x Genomics Cloud: Cell Ranger ARC Count v2.0.2**
   1. You should have 2 fastq file sets under one project: one for ATAC seq (I1, R1, R2, R3) and one for Gene Expression seq (I1, I2, R1, R2) data. ATAC library should be described under “Library of Feature Type” as “Single Cell Multiome ATAC” and Gene Expression/RNA seq library should be “Single Cell Multiome 3’ Gene Expression”. Start new analysis with both sets selected, Cell Ranger ARC Count v2.0.2 automatically pops up. Select start analysis with defaults selected. Results will be emailed within several hours (this sample took 2 hours, 44 min).
   2. Analysis 1:





The output files from CellRanger will be uploaded into R for further analysis