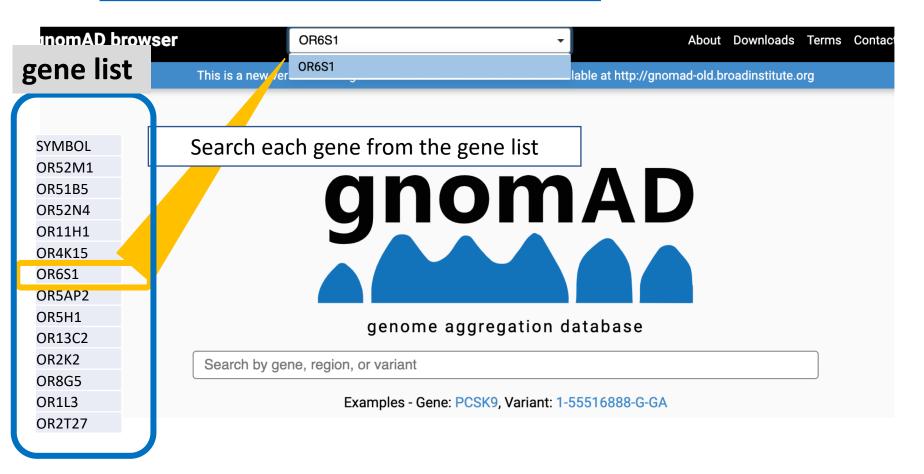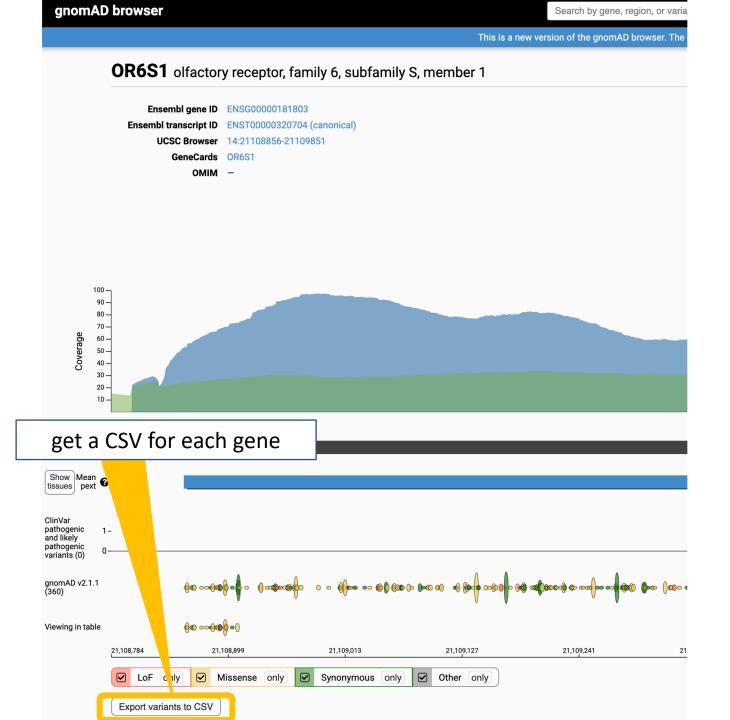# Retrieving data from gnomAD

Marie Saitou

Purpose

- We have a list of genes

- We would like to access an online database of genetic variation, and obtain the variation data of particular genes (web scraping?)

- They offer a downloadable files but it is 500Gb. So, working on the web seems to be faster
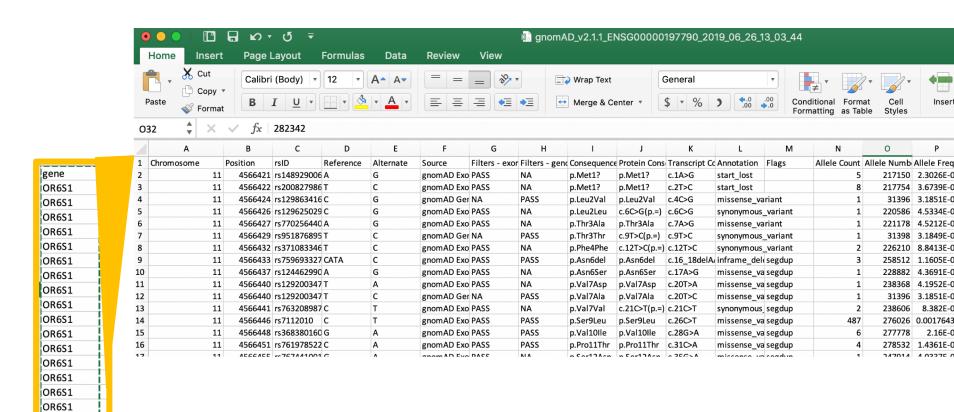
a database - gnomAD

- https://gnomad.broadinstitute.org/

gnomAD browser

Search by gene, region, or varia

This is a new version of the gnomAD browser. The

OR6S1 olfactory receptor, family 6, subfamily S, member 1

Ensembl gene ID	ENSG00000181803
Ensembl transcript ID	ENST00000320704 (canonical)
UCSC Browser	14:21108856-21109851
GeneCards	OR6S1
OMIM	—

get a CSV for each gene

Show tissues   Mean pext ❷

ClinVar pathogenic and likely pathogenic variants (0)

gnomAD v2.1.1 (360)

Viewing in table

21,108,784   21,108,899   21,109,013   21,109,127   21,109,241   21

☑ LoF only   ☑ Missense only   ☑ Synonymous only   ☑ Other only

Export variants to CSV

The output csv (genetic variation) head

I'd like to add a column of gene name

Merge the output files to a single csv

header
CSV of GENE1
CSV of GeNE2
CSV of GENE3

⋮