

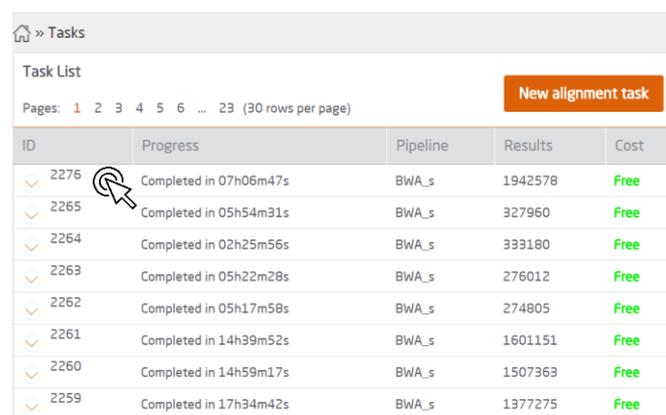
To solve Exercise III, you will need a genome browser. The Analyze Genomes platform offers a ready-to-use genome browser with several predefined, aligned genomic files which have been already processed for variant calling. You will be provided with a Task ID for the exercise. This manual guides you through the steps necessary to successfully complete this assignment.

Step 1. Logon into the Analyze Genomes Platform

In order to log on the Analyze Genomes Platform, please use the  button available on the course exercise page. It will automatically log you on the platform using your openHPI credentials.

Step 2. Choose alignment task

After login, you will be provided with a list of alignment tasks, as displayed below (Fig. 1). Simply click on the one indicated on the exercise description.

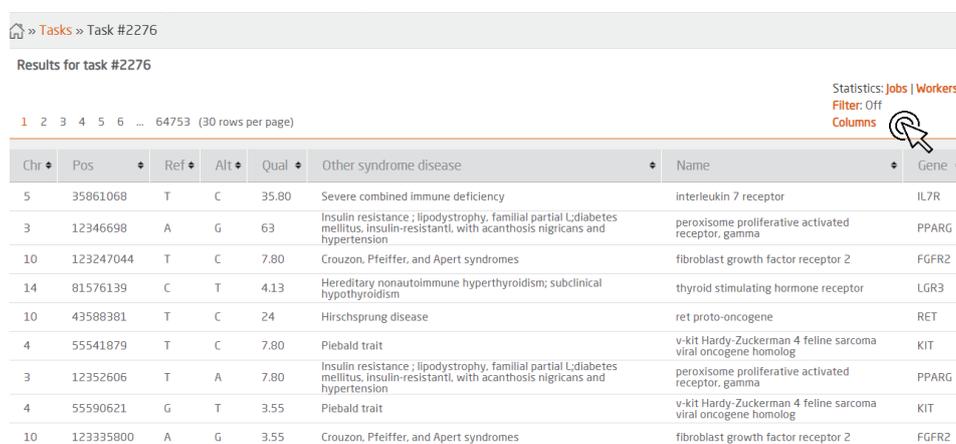


ID	Progress	Pipeline	Results	Cost
2276	Completed in 07h06m47s	BWA_s	1942578	Free
2265	Completed in 05h54m31s	BWA_s	327960	Free
2264	Completed in 02h25m56s	BWA_s	333180	Free
2263	Completed in 05h22m28s	BWA_s	276012	Free
2262	Completed in 05h17m58s	BWA_s	274805	Free
2261	Completed in 14h39m52s	BWA_s	1601151	Free
2260	Completed in 14h59m17s	BWA_s	1507363	Free
2259	Completed in 17h34m42s	BWA_s	1377275	Free

Figure 1. Task list

Step 3. Choose variant

Once you click on a given task, a list of genetic variants is displayed, according to the specific location it appears on the genetic code (Fig 2.). The last column on the left identifies the affected gene. Locate on the list the gene provided on the exercise and click on it to open the genome browser. **Important tip:** note that a click on any of the column headers (Chr, Pos, Ref, etc.) will sort the list alphabetically (A-Z, Z-A). That is helpful if you want to identify all variants related to a disease or gene. You can also access a given position changing the URL directly (see below) or by using the **Filter** options.



Chr	Pos	Ref	Alt	Qual	Other syndrome disease	Name	Gene
5	35861068	T	C	35.80	Severe combined immune deficiency	interleukin 7 receptor	IL7R
3	12346698	A	G	63	Insulin resistance ; lipodystrophy, familial partial ;diabetes mellitus, insulin-resistant, with acanthosis nigricans and hypertension	peroxisome proliferative activated receptor, gamma	PPARG
10	123247044	T	C	7.80	Crouzon, Pfeiffer, and Apert syndromes	fibroblast growth factor receptor 2	FGFR2
14	81576139	C	T	4.13	Hereditary nonautoimmune hyperthyroidism; subclinical hypothyroidism	thyroid stimulating hormone receptor	LGR3
10	43588381	T	C	24	Hirschsprung disease	ret proto-oncogene	RET
4	55541879	T	C	7.80	Piebald trait	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	KIT
3	12352606	T	A	7.80	Insulin resistance ; lipodystrophy, familial partial ;diabetes mellitus, insulin-resistant, with acanthosis nigricans and hypertension	peroxisome proliferative activated receptor, gamma	PPARG
4	55590621	G	T	3.55	Piebald trait	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	KIT
10	123335800	A	G	3.55	Crouzon, Pfeiffer, and Apert syndromes	fibroblast growth factor receptor 2	FGFR2

Figure 2. List of genetic variants

Step 4. Explore the genome browser

The questions asked on the exercise are based on the information provided on the genome browser. An example is shown on Fig. 3. The relevant genetic sequence is displayed, alongside with the respective translated aminoacids in both 5'-3' and 3'-5' direction. Note that the start codon (ATG), which codes for Methionine (M), is marked in dark grey. Similarly, end codons (i.e. TGA) are marked in red.

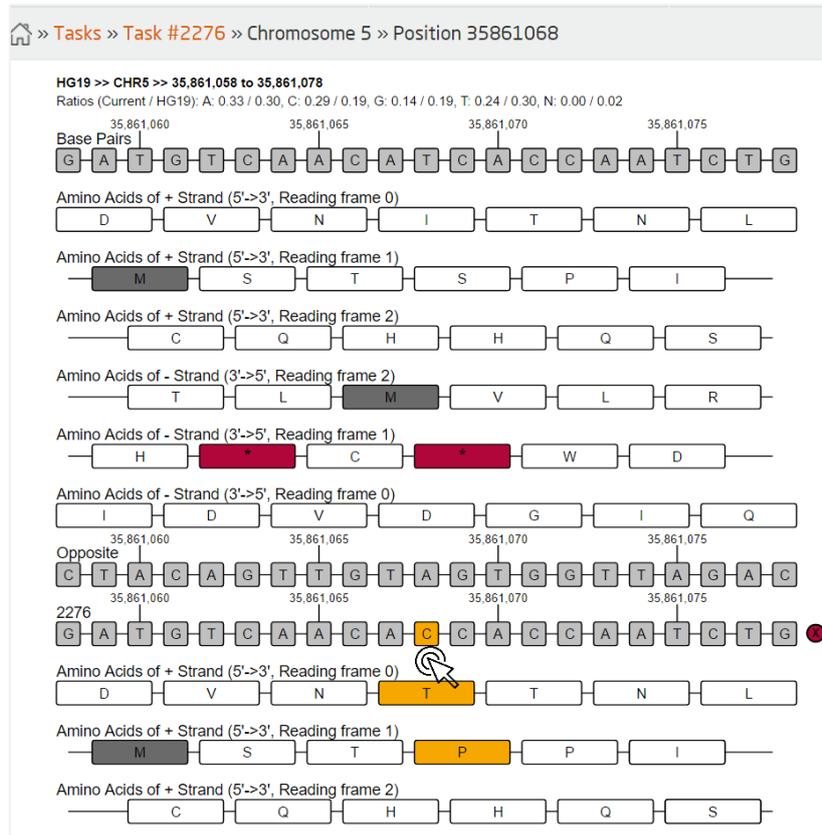


Figure 3. Genome browser

Step 5. Retrieve information from literature

With properly annotated genomic files, specific variants can be traced back to the studies that originally identified them for the first time and further references to it in available research. Click on the respective variant (cf. Fig. 3) to display more information. An infobox will be displayed on the right (Fig. 4). At this point, you're good to go! Just refer back to the genome browser to answer the questions. 😊

**MUTATION DETAILS FOR
CHR5 AT 35,861,068**

DGV

> Pubmed ID: [15286789](#)
> Short description: Twenty samples were obtained from normal individuals whose DNA are stored at the Coriell Cell Repository (Camden, New Jersey) with the following ethnic representation: 10 European/Caucasian, 4 Native American, 2 Chinese, 2 Indo-Pakistani, and 2 Sub-Saharan African. Twelve samples were from fixed cell pellets of cytogenetically normal individuals from the Brigham and Women's Cytogenetics Laboratory and 7 samples from cytogenetically normal individuals were obtained from the Hospital for Sick Children.

Figure 4. Literature information on a given genetic variant

Quickly finding a given position within the Genome Browser

If you have trouble finding a given position on the Genome Browser, you can quickly access it by changing the URL as follows:

<https://we.analyzegenomes.com/now/tasks/<<task id>>/<<chromosome number>>/<<position>>>

Just change <<task id>>, <<chromosome number>> and <<position>> accordingly and you will be directly taken to the position desired.