

[Wiki » R01 - r100 et intronSeeker sur données réelles »](#)

Méthodologie de sélection des échantillons depuis le site du NCBI

Dans Google, chercher "sequence set browser" -> <https://www.ncbi.nlm.nih.gov.ezproxy.u-pec.fr/Traces/wgs/?view=TLS>

<https://www.ncbi.nlm.nih.gov.ezproxy.u-pec.fr> > ... ▼ [Traduire cette page](#)

Sequence Set Browser :: NCBI

skip to main content. Facet Panel Facet Panel. External link. Please review our privacy policy.

Dans "Project type", sélectionner "TSA", puis préciser le nom de l'espèce recherchée dans "Term":

The screenshot shows the NCBI Sequence Set Browser interface. The search term is 'Acipenser'. The results table lists 10 projects. The first project is GETX01, which is a TSA project for Acipenser sturio. The table columns are: #, Prefix, Type, DIV, Organism, Biosample, Infospecific Name, Other Source, Contigs (Total Length, #), # in Prot, # Has Annot, # Update Date, # Create Date.

#	Prefix	Type	DIV	Organism	Biosample	Infospecific Name	Other Source	Contigs	# in Prot	# Has Annot	# Update Date	# Create Date
								\$ Total Length (Mbases)	\$ #			
1	GETX01	TSA	VRT	Acipenser sturio	PRJNA306186	SAMN0481347023	Issue, type: liver, spleen and kidney	58.9	91,687		2016-07-19	2016-07-19
2	GELU01	TSA	VRT	Acipenser oxyrinchus	PRJNA311275	SAMN0481348000 SAMN0481348001 SAMN0481348002 SAMN0481348003 SAMN0481348004 SAMN0481348005 SAMN0481348006 SAMN0481348007 SAMN0481348008 SAMN0481348009 SAMN0481348010 SAMN0481348011 SAMN0481348012 SAMN0481348013		196.7	179,564		2017-06-02	2017-06-02
3	GGQJ01	TSA	VRT	Acipenser sturio	PRJNA306040	SAMN056332004	strain: hybrid sturgeon	135.6	77,634		2019-01-15	2019-01-15
4	GGWJ01	TSA	VRT	Acipenser oxyrinchus	PRJNA400239	SAMN100031349 SAMN100031350 SAMN100031351 SAMN100031352 SAMN100031353 SAMN100031354 SAMN100031355 SAMN100031356 SAMN100031357 SAMN100031358 SAMN100031359 SAMN100031360 SAMN100031361 SAMN100031362 SAMN100031363 SAMN100031364 SAMN100031365 SAMN100031366 SAMN100031367		34.8	53,624		2018-09-25	2018-09-25

Chercher un projet avec un nombre de contigs suffisant (min 30 000), et avec plusieurs biosamples:

The screenshot shows the NCBI Sequence Set Browser interface. The search term is 'Drosophila'. The results table lists 15 projects. The first project is GADM01, which is a TSA project for Drosophila melanogaster. The table columns are: #, Prefix, Type, DIV, Organism, Biosample, Infospecific Name, Other Source, Contigs (Total Length, #), # in Prot, # Has Annot, # Update Date, # Create Date.

#	Prefix	Type	DIV	Organism	Biosample	Infospecific Name	Other Source	Contigs	# in Prot	# Has Annot	# Update Date	# Create Date
								\$ Total Length (Mbases)	\$ #			
1	GADM01	TSA	INV	Drosophila melanogaster	PRJNA172931			5.5	9,594		2014-01-02	2014-01-02
2	GADP01	TSA	INV	Drosophila erecta	PRJNA172931			7.4	15,456		2014-01-02	2014-01-02
3	GADQ01	TSA	INV	Drosophila melanogaster	PRJNA172931			10.3	18,361		2014-01-02	2014-01-02
4	GADR01	TSA	INV	Drosophila pseudoannulata	PRJNA172931			12.3	29,641		2014-01-02	2014-01-02
5	GADS01	TSA	INV	Drosophila melanogaster	PRJNA172931			10	19,353		2014-01-02	2014-01-02
6	GAEJ01	TSA	INV	Drosophila pseudoannulata	PRJNA172931		div. stage: adult issue, type: whole body	13.9	32,524		2014-01-02	2014-01-02
7	GARN01	TSA	INV	Drosophila serrata	PRJNA193659			10.1	11,328		2013-04-18	2013-04-18
8	GALP01	TSA	INV	Drosophila melanogaster	PRJNA208862			26.3	12,521		2014-06-03	2014-06-03
9	GBUQ01	TSA	INV	Trichopria drosophilae	PRJNA252238	SAMN02870433	div. stage: adult isolation source: reared on Drosophila melanogaster	23.9	19,090		2017-03-24	2017-03-22
10	GDUH01	TSA	ENV	insect metagenome	PRJNA295093	SAMN04075829 SAMN04075830 SAMN04075831 SAMN04075832	isolation source: Drosophila with parasite and bacterial infection	95.5	30,575		2015-10-28	2015-10-28
11	GECC01	TSA	INV	Drosophila melanogaster	PRJNA300939	SAMN02890087 SAMN02890088 SAMN02890089 SAMN02890090		14.8	31,233		2015-12-08	2015-12-08

Cliquer sur l'identifiant du projet:

Accession	Species	Strain	Library	Reads	Size	Date
3 GADQ01 TSA INV	<i>Drosophila melanogaster</i>	PRJNA172931		10.3	19,361	2014-01-02
4 GADQ01 TSA INV	<i>Drosophila pseudobscurus</i>	PRJNA172931		12.3	29,641	2014-01-02
5 GADQ01 TSA INV	<i>Drosophila melanogaster</i>	PRJNA172931		10	19,363	2014-01-02
6 GAEJ01 TSA INV	<i>Drosophila melanogaster</i>	PRJNA172931	dev. stage: adult tissue: whole body	13.9	32,524	2014-01-02
7 GAEJ01 TSA INV	<i>Drosophila melanogaster</i>	PRJNA193059		10.1	11,328	2013-04-18
8 GALP01 TSA INV	<i>Drosophila melanogaster</i>	PRJNA208852		26.3	12,521	2014-06-03
9 GRQJ01 TSA INV	<i>Trichopoda clavigera</i>	PRJNA252238	strain: wildtype dev. stage: adult isolation source: melanogaster tissue: whole organism	23.9	19,090	2017-03-24
10 GDUH01 TSA ENV	Insect metagenome	PRJNA295093	isolation source: <i>Drosophila</i> with parasite and bacterial infection	55.5	30,575	2015-10-28
11 GECM01 TSA INV	<i>Drosophila melanogaster</i>	PRJNA309399		14.8	31,233	2015-12-08
12 GH4H01 TSA INV	<i>Drosophila melanogaster</i>	PRJNA526247	Project is live tissue: Final abdominal tip	65.3	35,521	2020-06-19
13 GH4H01 TSA INV	<i>Drosophila melanogaster</i>	PRJNA526247	tissue: Final abdominal tip	46.7	40,037	2020-06-19
14 GH4H01 TSA INV	<i>Drosophila melanogaster</i>	PRJNA526247	tissue: Final abdominal tip	55.1	31,980	2020-04-07
15 GH4H01 TSA INV	<i>Drosophila melanogaster</i>	PRJNA526247	tissue: Final abdominal tip	40.8	31,071	2020-04-07

La description du projet s'affiche:

COVID-19 is an emerging, rapidly evolving situation.
Get the latest public health information from CDC: <https://www.cdc.gov/coronavirus>
Get the latest research information from NIH: <https://www.nih.gov/coronavirus>
Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>

ISA Brown Chickens Embryo HH26 Isoseq
Accession: PRJEB13248 ID: 324235

An embryo at HH stage 26 from ISA Brown chickens was collected and RNA was extracted. More...

Project Data

Resource Name	Number of Links
Sequence sets	
Nucleotide (total)	1,492,012
TSA-master	1
Transcript	1,492,012
SRA Experiments	33
Other samples	
BioSample	33

Related information

- BioSample
- Nucleotide
- Related Genes
- SRA
- Transcript
- TSA-master

Recent activity

- ISA Brown Chickens Embryo HH26 Isoseq (BioSample)
- ERAS06729 (33) (SRA)
- Understanding SRA Search Results - SRA Knowledge Base
- BioSample for BioProject (Select 324235) (33) (BioSample)
- ERAS06729 NOT cluster_09seq (33) (SRA)

Récupérer l'"accession number":

ISA Brown Chickens Embryo HH26 Isoseq
Accession: PRJEB13248 ID: 324235

An embryo at HH stage 26 from ISA Brown chickens was collected and RNA was extracted. More...

Project Data

Resource Name	Number of Links
Sequence sets	
Nucleotide (total)	1,492,012
TSA-master	1
Transcript	1,492,012
SRA Experiments	33
Other samples	
BioSample	33

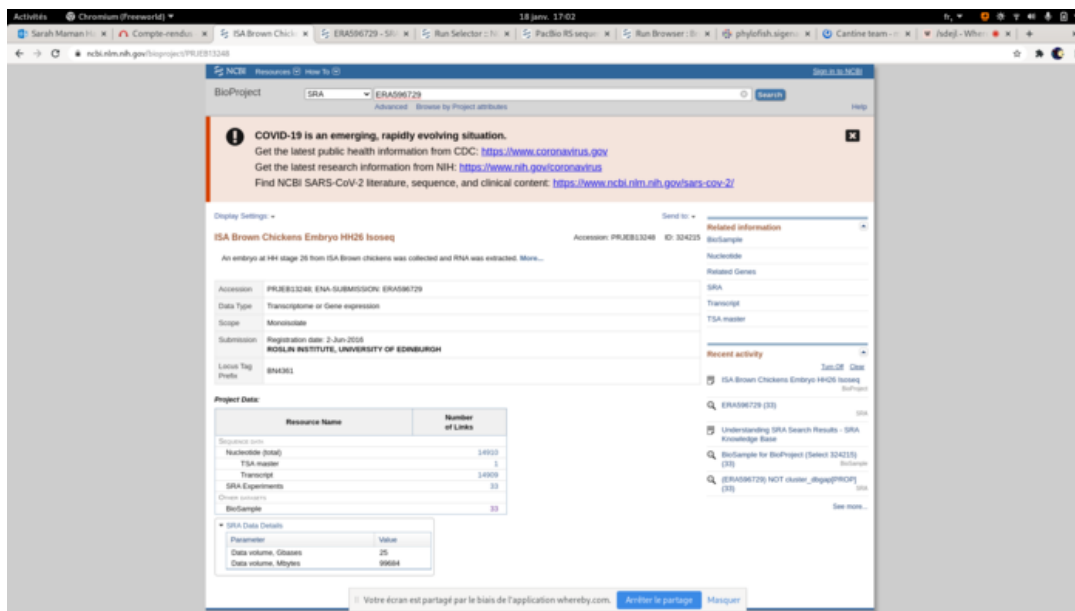
Related information

- BioSample
- Nucleotide
- Related Genes
- SRA
- Transcript
- TSA-master

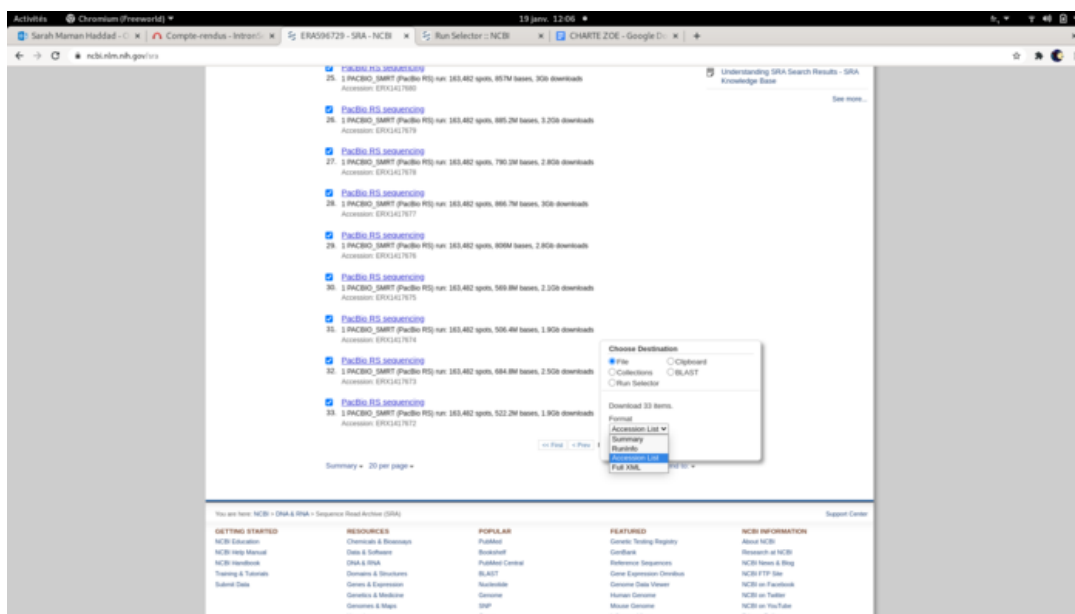
Recent activity

- ISA Brown Chickens Embryo HH26 Isoseq (BioSample)
- ERAS06729 (33) (SRA)
- Understanding SRA Search Results - SRA Knowledge Base
- BioSample for BioProject (Select 324235) (33) (BioSample)
- ERAS06729 NOT cluster_09seq (33) (SRA)

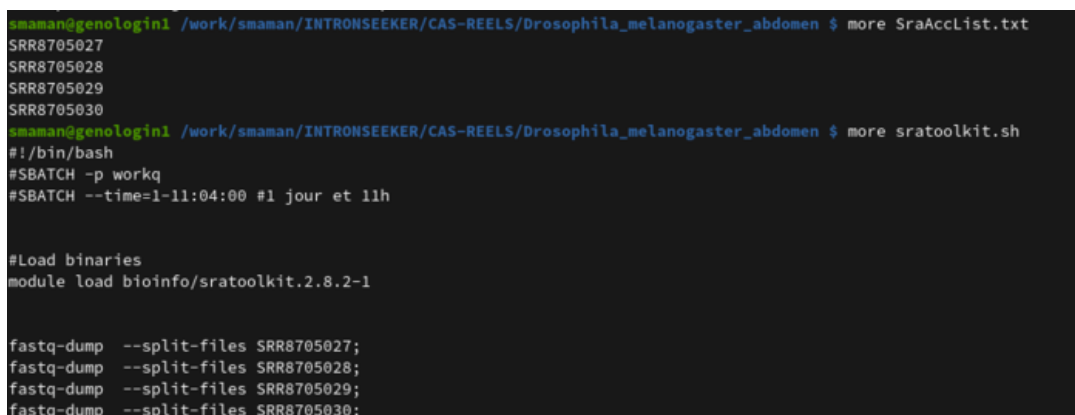
Rechercher l'accession number dans SRA:



Sélectionner une liste de FASTQ ILLUMINA seulement puis récupérer cette liste dans un fichier "Accession List"



Télécharger les FASTQ dans son work à l'aide du script sratoolkit.sh :



0-sequence_set_browser.png (19,795 ko) Sarah Maman, 11/10/2021 14:16

0-term_TSA_search.png (279,235 ko) Sarah Maman, 11/10/2021 14:16

2-cliquer_sur_identifiant_du_projet.png (286,642 ko) Sarah Maman, 11/10/2021 14:16

7-download-fastqR1R2_dans_work_avec_sratoolkit.png (46,053 ko) Sarah Maman, 11/10/2021 14:16

1-chercher_un_projet_pour_un_animal.png (324,525 ko) Sarah Maman, 11/10/2021 14:16

6-Selectionner_la_liste_des_fastq_a_recup.png (237,86 ko) Sarah Maman, 11/10/2021 14:16

run_selector.png (254,755 ko) Sarah Maman, 11/10/2021 14:16

5-rechercher_dans_sra_l-accession_number-en_haut_de_la_page.png (221,049 ko) Sarah Maman, 11/10/2021 14:16

[4-recuperer_accession_number.png](#) (219,178 ko)  Sarah Maman, 11/10/2021 14:16

[3-affichier_description_du_projet.png](#) (219,192 ko)  Sarah Maman, 11/10/2021 14:16