Prokaryote - Tutorial Hp 26695

This introductory section provides an overview of **Prokaryote** pipeline drafting and design. The vertical gray rectangles correspond to the website sections.



This document shows how to perform all the steps of an example analysis with the data provided for Prokaryote pipeline. For further information about the forms and the interpretation of the results, refer to the User Guide.

Input Files: dataset Prokaryote_Hp26695.zip

Download a zip folder containing the raw datasets from this link (1.1 GB).

Mandatory inputs for InteractomeSeq - Prokaryote Hp 26695 execution are:

- genome reference file in FASTA format (either a custom annotation file or one selected from the drop-down menu) (Organism).
 - GCF_000008525.1_ASM852v1_genomic.fna (1.6 MB)
- a genome annotation (either a custom annotation file or one selected from the drop-down menu) (Organism).
 - GCF_000008525.1_ASM852v1_genomic.gff (620 KB)
- Raw Data files, FASTA or FASTQ format for query reads are allowed in the input, therefore the web interface additionally allows the submission of compressed files (gz format) to reduce the time of data upload (DataSets).
 - HP_genomic_26695_R1.fastq.gz (97 MB)
 - HP_genomic_26695_R2.fastq.gz (133 MB)
 - Sel_AG_R1.fastq.gz (243 MB)
 - Sel_AG_R2.fastq.gz (335 MB)
 - Sel_HPneg_R1.fastq.gz (66 MB)
 - Sel_HPneg_R2.fastq.gz (91 MB)
 - Sel_HPpos_R1.fastq.gz (74 MB)
 - Sel_HPpos_R2.fastq.gz (104 MB)

InteractomeSeq requires the user to upload at least two datasets. The input datasets must be generated with the same sequencing platform.

Time and memory usage for this tutorial

Analysis Step	File	Excution time	Output size
			46.140
Mapping	26695_\$5	~ 9m	46 MB
Mapping	HpNegativeControl	~ 6m	8 MB
Mapping	HpPositiveControl	~ 12m	13 MB
Mapping	AtrophicGastritis	~ 28m	28 MB
Domain Definition	26695_S5	~ 2m	357 KB
Domain Definition	HpNegativeControl	~ 35s	84 KB
Domain Definition	HpPositiveControl	~ 50s	126 KB
Domain Definition	AtrophicGastritis	~ 1m	222 KB
Domain Enrichment	26695_S5 - HpNegativeControl	~ 20s	30 KB
Domain Enrichment	26695_S5 - HpPositiveControl	~ 25s	39 KB
Domain Enrichment	26695_S5 - AtrophicGastritis	~ 30s	82 KB
Domain Subtraction	HpPositiveControl - AtrophicGastritis	< 5s	52 KB
Domain Subtraction	HpNegativeControl - HpPositiveControl	< 5s	27 KB
Domain Subtraction	HpNegativeControl - AtrophicGastritis	< 5s	67 KB
Domain Intersection	HpPositiveControl - AtrophicGastritis	< 5s	97 KB
Domain Intersection	HpNegativeControl - HpPositiveControl - AtrophicGastritis	< 5s	122 KB

Create a project

Click the button **START** to create a new project. Give the project a name (mandatory), a description (optional) and an email address (to receive messages during the execution. Select **PROKARIOTE** as project type and **SAVE**.

Create a Projec	t
Project Name 🛈	Hp 26695
Project Description	Helicobacter pylori 26695 Project
Email Address 🛛	my@email.com
Project Type 0	EUKARYOTE PROKARYOTE
	SAVE 4 CANCEL 🗶

The project now has an ID that is a link, active for 15 days. Save this link if you didn't provide an email address (the link will be sent via email).

Information S	ummary		
C			
Project Name	Hp 26695		
Project ID	yb3qjkao8p40f6rbtb0mywsg5p7w %		
Project Description	Helicobacter pylori 26695 Project		
Project Type	Prokaryote		
Project Status			
Creation Date	Monday March 9, 2020 - 19:03:09	Expiration Date	Tuesday March 24, 2020 - 19:03:09

Uploading

Clik on UPLOADING to upload the FASTQ files.

Select the **Custom Annotation** form to upload the genome files. Upload the reference file (**GCF_000008525.1_ASM852v1_genomic.fna**), and press **VALIDATE**. After the validation, a preview of the genomic sequences will be shown in a table.

Annotation				
Reference				
Reference File O SELECT FILE			Drop File	
File	Size	Date	Progress	Status Actions
GCF_000008525.1_ASM852v1_genomic.fna	📩 1.61 M	IB 12/20/2019 - 14:36:59	100%	REMOVE
	PR	REVIEW 👁 VALIDATE 🗹		
Reference File - Valid				
# REFERENCES : 1				
Header			Sequence	
NC_000915.1 Helicobacter pylori 26695 chromosome, comp	olete genome T	GATTAGTGATTAGTGATTAGTGATTA	AGTGATTAGTGATTAGTGATTAGTGATTAGTGA	TTAGTGATTAGTGATTAGTGA
				5 10 25

Upload the annotation (**GCF_000008525.1_ASM852v1_genomic.gff**), select the annotation format (**GFF**) and press **VALIDATE** (it is not necessary to compile the other fields, since GFF is a standard format).

	e 🛈 SELEC	TFILE			Drop File		
	File		Size	Date	Progress	Status	Actions
GCF_000008525.1_ASI	M852v1_genomic.ç	yff 🕹	0.61 MB 12/	20/2019 - 14:39:43	100%	~	TREMOVE
File Format	Ð BET) GFF	CSV/TSV	Head # Hoad	der Line 🛛 💭		
Column Separator	TAB	▼.					
			PREVIEW @	VALIDATE 🗹			
Annotation File - \	/alid						
# ANNOTATIONS : 1	,						
# ANNOTATIONS : 1 Chromosome	Start	End Stran	d Locus Tag	Gene Name	Descr	ription	
# ANNOTATIONS : 1 Chromosome NC_000915.1	Start 217	End Stram	Id Locus Tag	Gene Name	Descr transcription antitermination pro	iption tein NusB	
# ANNOTATIONS : 1 Chromosome NC_000915.1 NC_000915.1	Start 217 635	End Stran 633 - 1105 -	HP0001 HP0002	Gene Name nusB ribH	Descr transcription antitermination pro 6%2C7-dimethyl-8-ribityllumazin	tein NusB e synthase	
# ANNOTATIONS: 1 Chromosome NC_000915.1 NC_000915.1 NC_000915.1	Start 217 635 1115	End Stran 633 - 1105 - 1945 -	HP0001 HP0002 HP0003	Gene Name nusB ribH	Descr transcription antitermination pro 6%2C7-dimethyl-8-ribityllumazin 2-dehydro-3-deoxyphosphooctor	tein NusB e synthase nate aldolase	
# ANNOTATIONS : 1 Chromosome NC_000915.1 NC_000915.1 NC_000915.1 NC_000915.1	Start 217 635 1115 1932 1932	End Stran 633 - 1105 - 1945 - 2597 -	d Locus Tag HP0001 HP0002 HP0003 HP0004	Gene Name nusB ribH	transcription antitermination pro 6%2C7-dimethyl-8-ribityllumazin 2-dehydro-3-deoxyphosphooctor carbonic anhydrase IcfA	tein NusB e synthase nate aldolase	

In the **DataSets** section, click on **SELECT FILE** or drag&drop files to upload FASTQ datasets:

- HP_genomic_26695_R1.fastq.gz
- HP_genomic_26695_R2.fastq.gz
- Sel_HPneg_R1.fastq.gz
- Sel_HPneg_R2.fastq.gz
- Sel_HPpos_R1.fastq.gz
- Sel_HPpos_R2.fastq.gz
- Sel_AG_R1.fastq.gz
- Sel_AG_R2.fastq.gz

	DataSets	S Files STQ Format SELECT FILE Drop File		
Ē	Raw Data Files		Drop File	
	FASTA/FASTQ Format 🖲	SELECT FILE	Drop File	
L				

Dataset loading visualization:

aw Data Files						
FASTA/FASTQ Format @ SELECT FILE				Drop File		
File		Size	Date	Progress	Status	Actions
HP_genomic_26695_R1.fastq.gz	*	96.64 MB	12/20/2019 - 17:34:28	100%	~	TREMOVE
HP_genomic_26695_R2.fastq.gz	*	133.02 MB	12/20/2019 - 17:34:51	100%	~	TREMOVE
Sel_AG_R1.fastq.gz	*	243.40 MB	12/20/2019 - 17:35:14	100%	~	T REMOVE
Sel_AG_R2.fastq.gz	*	334.96 MB	12/20/2019 - 17:36:33	100%	~	TREMOVE
Sel_HPneg_R1.fastq.gz	*	66.21 MB	12/20/2019 - 17:36:39	100%	×	TREMOVE
Sel_HPneg_R2.fastq.gz	*	90.75 MB	12/20/2019 - 17:36:48	100%	×	TREMOVE
Sel_HPpos_R1.fastq.gz	¥	74.03 MB	12/20/2019 - 17:36:55	100%	×	TREMOVE
Sel HPpos R2.fasto.gz	±	104.08 MB		42%	0	X CANCEL

Mapping

In the **MAPPING** section, by clicking on the button + **MAPPING**, 4 sub-sections will appear on the screen:

1. **Mapping Params.** Select paired-end reads (in this example: HP_genomic_26695_R1.fastq.gz and HP_genomic_26695_R2.fastq.gz)

	Mapping List							
	+ MAPPING 2	TOTAL : 0					E COLU	MNS 🗕
	Info	Label	÷	Status 🌲	Date 👙	Log Ou	tput Edit	Delete
-								
::::	Mapping :: Insert							
	Mapping Params							
	Sequencing Type 🛛	SINGLE-READ	PAIRED-ENDS					
	Read Forward File 🚯			Read Reverse File 0		D0 6		- 1
		HP_genomic_20095_K1.tastq.gz	•	_	HP_genomic_26695_	RZ.1851Q.92		
	Mapping Label	26695_S5						

2. **Organism.** Preloaded FASTA file that will be used as reference to align the sequences.

🚍 Organism		
Reference 🕄	GCF_000008525.1_ASM852v1_genomic.fna	
Annotation	GCF_000008525.1_ASM852v1_genomic.gff	

- 3. Adapters. Select Custom Adapters and insert:
 - a. Forward Read 5' Adapter: GCAGCAAGCGGCGCGCATGCCACTAGTGGGAT
 - b. Forward Read 3' Adapter: CCCAGAGCAA
 - c. Reverse Read 5' Adapter: GGGATTGGTTTGCCGCTAGCGGAGAT
 - d. Reverse Read 3' Adapter: CCCAGAGCAA

III Adapters	
Adapters	Custom Adapters 💌
Forward Read 5' Adapter 🕄	GCAGCAAGCGGCGCGCACTAGTGGGAT
Forward Read 3' Adapter 🚯	CCCAGAGCAA
Reverse Read 5' Adapter 0	GGGATTGGTTTGCCGCTAGCGGAGAT
Reverse Read 3' Adapter	CCCAGAGCAA

4. **Trimming Params.** Selection of minimum length of sequence and number of mismatch allows, reads below this threshold will be discarded. Use the default, and click **EXECUTE**

₽	Trimming Params					
	Min Clone Length	100 🖨	 •	Allowed Mismatches	3 🛓	
			EXECUTE 4 RESE	T 🛍 🛛 CANCEL 🗙		

This mapping will appear in the list of running mappings.

Mapping Li	st									
+ MAPPING	C TOTAL:								🖽 COLU	MNS 🔻
Info	Label	÷	Status	÷	Date	÷	Log	Output	Edit	Delete
	26695_S5		Running		20/12/2019 - 17:	47:09	۲		Ø	×
								5	i 10	25
									_	

Here are the parameters for the other 3 mappings in this example:

1. Sel_HPneg_R1

SequencingType Paired-Ends

Forward Read File Sel_HPneg_R1.fastq.gz

Reverse Read File Sel_HPneg_R2.fastq.gz

Reference GCF_000008525.1_ASM852v1_genomic.fna

Annotation GCF_000008525.1_ASM852v1_genomic.gff

AdapterType Custom

Forward Read 5' Adapter GCAGCAAGCGGCGCGCATGCCACTAGTGGGAT

Forward Read 3' Adapter CCCAGAGCAA

Reverse Read 5' Adapter GGGATTGGTTTGCCGCTAGCGGAGAT

Reverse Read 3' Adapter CCCAGAGCAA

Min Clone Length 100

Allowed Mismatches 3

2. Sel_HPpos_R1

SequencingType Paired-Ends

Forward Read File Sel_HPpos_R1.fastq.gz

Reverse Read File Sel_HPpos_R2.fastq.gz

Reference GCF_000008525.1_ASM852v1_genomic.fna

Annotation GCF_000008525.1_ASM852v1_genomic.gff

AdapterType Custom

Forward Read 5' Adapter GCAGCAAGCGGCGCGCATGCCACTAGTGGGAT

Forward Read 3' Adapter CCCAGAGCAA

Reverse Read 5' Adapter GGGATTGGTTTGCCGCTAGCGGAGAT

Reverse Read 3' Adapter CCCAGAGCAA

Min Clone Length 100

Allowed Mismatches 3

3. Sel_AG_R1

SequencingType Paired-Ends

Forward Read File Sel_AG_R1.fastq.gz Reverse Read File Sel_AG_R2.fastq.gz Reference GCF_000008525.1_ASM852v1_genomic.fna Annotation GCF_000008525.1_ASM852v1_genomic.gff AdapterType Custom Forward Read 5' Adapter GCAGCAAGCGGCGCGCATGCCACTAGTGGGAT Forward Read 3' Adapter GCAGCAAGCGACAA Reverse Read 3' Adapter CCCAGAGCAA Reverse Read 3' Adapter CCCAGAGCAA Min Clone Length 100 Allowed Mismatches 3

This is the list of running mappings:

Mapping List							
TOTAL : 4						E COL	umns 🗸
Info	Label	÷	Status	÷	Date 🔶	Log	Output
×	26695_S5		Done		01/10/2019 - 14:00:00	۲	
~	HpNegativeControl		Running		01/10/2019 - 14:00:00	•	
~	HpPositiveControl		Running		01/10/2019 - 14:00:00	۰	
~	AtrophicGastritis		Running		01/10/2019 - 14:00:00	۰	==

Domain Analysis

Domain analysis is composed by four sheets:

- 1. Domain Definition
- 2. Domain Enrichment
- 3. Domain Subtraction
- 4. Domain Intersection

Domain Definition takes as input the mapping file previously generated. Define the domains for each of the 4 datasets uploaded and mapped by choosing them in the **Mapping** menu, then press **Execute**.

Domain Definition :	:: Insert
E Domain Definition Para	ams
Mapping 🛛	26695_S5 🗨
Domain Definition Label 🛛	26695_S5
a Organism	
Reference ()	GCF_000008525.1_ASM852v1_genomic.fna
Annotation O	GCF_000008525.1_ASM852v1_genomic.gff
I	
	EXECUTE 🐐 RESET 🛍 CANCEL 🗙

Domain Definition List shows the running domain definitions

Domain Definition	on List									
+ DOMAIN DEFINITION	TOTAL: 4								🖽 COLU	MNS 👻
Info	Label	÷	Status	÷	Date	÷	Log	Output	Edit	Delete
	26695_S5		Running		23/12/2019 - 9:49:	33	•		Ø	×
~	HpNegativeControl		Running		23/12/2019 - 9:49:	40	۲		œ	×
~	HpPositiveControl		Running		23/12/2019 - 9:49:	46	۲		Ø	×
v	AtrophicGastritis		Running		23/12/2019 - 9:49:	51	٢		Ø	×

Info – Drop-down menu with information of Mapping input file.

Label – Sample label.

Status – When the execution ends successfully, the button turns green, otherwise, it turns red.

Date – Day and time of analysis execution

Log – Button that hides/opens a box with execution log file. When the execution is running, the log shows the possibility to stop it:

🇱 Domain Defini	tion :: Log :: 2669	95_\$5	
I STATUS ♥	STOP 🗙	Domain Definition Running O Processing	
Computing the depth-o	f-coverage complete.		
		CLOSE 🗙	

When the process has finished, the log shows summary information:

Domain Definition :: Log :: 2669	5_S5
≣ STATUS 📀	Domain Definition Done 🗹 Completed Processing
Conversion Tabular ==> FastA complete.	
Representative clones picking complete.	
Blastn analysis complete.	
Output Bed parsed complete.	
Checking clones gene intersection complete.	
Filtering Clones outside CDS.	
Clone cluster count.Complete	
Filtering cluster complete.	
Filtering cluster complete.	
Pybedtools getfasta complete.	
Checking clones gene intersection complete.	
Add gene description complete.	
Conversion FastA ==> Tabular complete.	
Add nucleotide sequence complete.	

Output - Hides/opens panel with output preview

Domain Def	finition Output File			4	DOWNLOAD					
TOTAL :	2,986						RESET FILTERING	\downarrow_Z^A RESET SORTING	⊞c	OLUMNS 🔻
Info	Chromosome	÷	Clone Start 👙	Clone End 💠	Clone Length 🛛 🌩	Gene Start 💠	Gene End 💠	Gene	\$	Strand \$
	NC_000915.1		346	526	180	217	633	HP0001		-
	NC_000915.1		724	1073	349	635	1105	HP0002		-
~	NC_000915.1		1178	1721	543	1115	1945	HP0003		-
~	NC_000915.1		1775	1955	170	1115	1945	HP0003		-
	NC_000915.1		1983	2463	480	1932	2597	HP0004		-
~	NC_000915.1		2751	3018	267	2719	3402	HP0005		+
	NC_000915.1		3060	3168	108	2719	3402	HP0005		+
~	NC_000915.1		3789	3897	108	3403	4233	HP0006		+
~	NC_000915.1		3900	4035	135	3403	4233	HP0006		+
	NC_000915.1		5743	5985	242	5241	7145	HP0009		-

Genome Browser 🔄 JBROWSE		
Available Tracks	Genome Track View Help	co Share
X filter tracks	0 200,000 400,000 500,000 800,000 1,000,000 1,200,000 1,400,000	1,600,000
▼ Annotation 2		2
Reference Sequence	550,000 5555,000	56
NC_000915.1 Annotation	© Reference Sequence Zoom in to see sequence Zoom in to see sequence	.e
▼ Domain Definition 4		
✓ 26695_S5	© NC_000915.1 Annotation NP_207319.1 NP_207321.1 NP_207323.1	
HpNegativeControl	15.1 NP_207317.1 NP_207320.1 NP_207322.1	
AtrophicGastritis	NP_207318.1	
	8	
	© 26695_S5 7-	

Domain Enrichment takes as input the Genomic and Target output of Domain Definition step. Enrich domains for each defined domain by selecting the Genomic and the target domains. In the current example, select:

- 26695_S5 and NegativeControl
- H26695_S5 and PositiveControl
- 26695_S5 and AtrophicGastritis

Jomain Enrichmen	t :: Insert
Genomic Domain Definition	26695_S5
Target Domain Definition ()	HealthyControl
Domain Enrichment Label Ø	26695_S5 + PositiveControl
	EXECUTE 🗲 RESET 🛍 CANCEL 🗙

Domain Enrichment List

+ DOMAIN ENRICHMENT	C TOTAL : 3							III COLU	MNS 👻
nfo	Label	÷	Status	÷	Date 🔶	Log	Output	Edit	Delete
	26695_S5 - HpNegativeControl		Done		23/12/2019 - 10:49:18	•		Ø	×
v	26695_S5 - HpPositiveControl		Running		23/12/2019 - 10:49:34	•		œ	×
v	26695_S5 - AtrophicGastritis		Running		23/12/2019 - 10:49:47	۲		C2	×

Info – Drop-down menu with information of Domain Definition input file.

Label – Sample label.

Status – When the execution ends successfully, the button turns green, otherwise, it turns red.

Date - Day and time of analysis execution

Log – Button that hides/opens a box with execution log file. When the execution is running, the log shows the possibility to stop it, otherwise the log shows summary information:

Domain Enrichment :: Lo	g :: 26695_S5 + HealthyControl
≣ STATUS 🛇	Domain Enrichment Done 🕢 Completed Processing
Parsing of mapping output file complete.	
Parsing of mapping output file complete.	
Parsing of domain definition output file con	aplete.
Bedtools coverage complete.	
Bedtools coverage complete.	
Parsing output bedtools coverage complete	<u>a</u> .
Parsing output bedtools coverage complete	a.
Differential expression analysis complete.	

Output - Hides/opens panel with output preview

TOTAL	275						RESET FILTERIN	G ↓ RES	ET SORTING	E COLUMNS
Info	Chromosome 💠	Clone Start \$	Clone End 💠	Clone Length \$	Gene Start 💠	Gene End 💠	Gene 🌩	Strand \$	Log FC \$	Adjust PValue [‡]
	NC_000915.1	238	373	135	217	633	HP0001	-	2.4653	3.0548e-2
	NC_000915.1	8395	8569	174	7603	9243	HP0010		2.1790	3.4764e-3
	NC_000915.1	10861	11046	185	9911	11590	HP0012	+	1.8962	1.3982e-2
	NC_000915.1	14979	15288	309	14248	16611	HP0017	+	3.3727	7.6874e-5
	NC_000915.1	16787	17147	284	16863	18272	HP0018	+	2.0296	6.5923e-3
	NC_000915.1	17966	18052	86	16863	18272	HP0018	+	6.7736	7.9131e-15
~	NC_000915.1	33910	33988	78	32680	34905	HP0033	+	2.9902	9.8855e-5
~	NC_000915.1	41903	42052	149	40651	42063	HP0043	+	5.7755	1.4830e-4
~	NC_000915.1	43269	43360	91	43243	44175	HP0045	+	2.5775	5.2191e-3
~	NC_000915.1	46430	46492	62	46042	48351	HP0048	-	2.8890	3.6349e-3
«	1 2 3	. 28 »							5 10	25 50

Domain Subtraction takes as input two differentially enriched epitopes/domains lists, one defined as Control Domain Enrichment and one defined as Selection Domain Enrichment. In this example, select the enrichments defined by the couples:

- 26695_S5 HpPositiveControl
- 26695_S5 AtrophicGastritis

then

- 26695_S5 HpNegativeControl
- 26695_S5 HpPositiveControl

and

- 26695_S5 HpNegativeControl
- 26695_S5 AtrophicGastritis

Control Domain Enrichment	26695_S5 + NegativeControl
Selection Domain Enrichment ()	26695_S5 + PositiveControl
Domain Subtraction Label 🚯	HpNegativeControl - HpPositiveControl
Params	
Overlap 🕄	0.5

Domain subtraction list:

+ DOMAIN SUBTRACTION 2 TOTAL: 3										
nfo	Label	÷	Status	÷	Date	÷	Log	Output	Edit	Delete
~	HpPositiveControl - AtrophicGastritis		Done		06/03/2020 - 17	:28:22	۲		Ø	×
~	HpNegativeControl - HpPositiveControl		Done		06/03/2020 - 17	57:53	•		ß	×
V	HpNegativeControl - AtrophicGastritis		Done		06/03/2020 - 17	58:37	۲		C2	×

Info – Drop-down menu with information of Domain Enrichment input file.

Label – Sample label.

Status – When the execution ends successfully, the button turns green, otherwise, it turns red.

Date – Day and time of analysis execution

Log – Button that hides/opens a box with execution log file. When the execution is running, the log shows the possibility to stop it, otherwise the log shows summary information:

Domain Subtraction :: Log :: HpNegativeControl - HpPositiveControl							
≣ STATUS 🛇	Domain Subtraction Done 🐨 Completed Processing						
Substraction domains complete.							

Output – Hides,	opens panel/	l with output	preview
------------------------	--------------	---------------	---------

TOTAL: 251 ▼ RESET FILTERING LAR RESET SORTING E COLUMNS -										
Info	Chromosome 💠	Clone Start ≑	Clone End 💠	Clone Length ÷	Gene Start 👙	Gene End 💠	Gene 🌩	Strand \$	Log FC 🌩	Adjust PValue 🍦
	NC 000915.1	5873	6027	154	5241	7145	HP0009	-	2.7581	1.3126e-3
	NC_000915.1	6284	6387	103	5241	7145	HP0009	-	5.1679	3.4435e-3
	NC_000915.1	8695	8837	142	7603	9243	HP0010	-	1.7501	2.7446e-2
	NC_000915.1	11762	12001	239	11587	12639	HP0013	+	1.7709	4.1700e-2
	NC_000915.1	13992	14096	104	13983	14246	HP0016	+	1.8357	4.3285e-2
	NC_000915.1	21880	21964	84	21152	22717	HP0022	-	2.2877	5.3086e-3
	NC_000915.1	26388	26480	92	26078	27358	HP0026	-	2.1189	8.6137e-3
	NC_000915.1	36527	36705	149	36556	37611	HP0037	+	3.0844	2.6988e-3
	NC_000915.1	37340	37464	124	36556	37611	HP0037	+	2.8550	1.2769e-3
~	NC_000915.1	60677	60803	126	57741	61298	HP0056	-	2.3361	1.1336e-2
¢	1 2 3	26 »							5 10	25 50

Domain Intersection takes as input two differentially enriched epitopes/domains lists output of Domain Definition step. Domain Intersection allows one to screen for overlaps between two sets of epitopes/domains lists. Select both the subtractions previously executed and press **EXECUTE**

Domain Intersection	on :: Insert
Selections ()	HpPositiveControl - AtrophicGastritis
	HpNegativeControl - HpPositiveControl
	HpNegativeControl - AtrophicGastritis
Domain Intersection Label Ø	HpPositiveControl - AtrophicGastritis
	EXECUTE 1/ RESET 1 CANCEL 🗙

Domain intersection list:

Domain Intersection List										
+ DOMAIN INTER	RSECTION								III COLU	MNS 🔻
Info	Label	÷	Status	÷	Date	÷	Log	Output	Edit	Delete
	HpPositiveControl - AtrophicGastritis				09/03/2020 - 11	12:35	۰		Ø	×
									5 10	25

Info – Drop-down menu with information of Domain Enrichment input file.

Label – Sample label.

Status – When the execution ends successfully, the button turns green, otherwise, it turns red.

Date - Day and time of analysis execution

Log – Button that hides/opens a box with execution log file. When the execution is running, the log shows the possibility to stop it, otherwise the log shows summary information:

Domain Intersection :: Log :: HpPositiveControl - AtrophicGastritis							
I STATUS ♥	Domain Intersection Done 🗹 Completed Processing						
Intersection domains complete.							

Output - Hides/opens panel with output preview



Results

The summary of all the analysis performed is showed in the results page

C TOTAL: 5	
nalvsis 🗢	
Label	Date
a Mapping	
HealthyControl	27/02/2020 - 11:46:51
PositiveControl	27/02/2020 - 11:55:37 📥
26695_S5	27/02/2020 - 15:45:00
Atrophic_Gastritis	27/02/2020 - 11:55:40
Domain Definition	
26695_S5	09/03/2020 - 18:52:32
HealthyControl	27/02/2020 - 12:30:08
PositiveControl	27/02/2020 - 12:30:13
Atrophic_Gastritis	27/02/2020 - 12:30:16
Domain Enrichment	_
26695_S5 + HealthyControl	27/02/2020 - 12:50:59
26695_S5 + PositiveControl	27/02/2020 - 12:51:01
26695_S5 + Atrophic_Gastritis	27/02/2020 - 12:51:04
Domain Subtraction	_
HpPositiveControl - AtrophicGastritis	06/03/2020 - 17:28:22
HpNegativeControl - HpPositiveControl	06/03/2020 - 17:57:53
HpNegativeControl - AtrophicGastritis	06/03/2020 - 17:58:37
Domain Intersection	
HpNegativeControl - HpPositiveControl - AtrophicGastritis	09/03/2020 - 13:51:27
HpPositiveControl - AtrophicGastritis	09/03/2020 - 11:12:35

And a genome browser with the results aligned to the reference genome

Genome Browser 🔛 JBROWSE								
Available Tracks	Genome Track	View Help						co Share
X filter tracks	0 200,000	400,000	600,000	800,000	1,000,000	1,200,000	1,400,000	1,600,000
✓ Annotation 2		🗩 Q 0	• •	NC_000915.1 -	NC_000915.1:66	671001000672 (3	33.57 K Go 🔜 🧧	S
Reference Sequence		750,000			875,000			1,000.
Custom Annotation	Reference Sequence		Zoom ir	n to see sequence		Zoom ir	n to see sequence	
✓ Domain Definition 4								
	Custom Annotation p	er 5,000 bp		10				
HealthyControl PoritiveControl								
Atrophic_Gastritis								
				8 -				
	© 26695_S5	Las de la		7				a beau
	haddittelid, hat telti	al date in the second second		a al la cuá Étala l	ala tali a di sila di di	that kining t	ahr Mathalia Iai	All date
						ULI UNI		
					الالكة التكري		الكالعدانية والع	
	Unable Control			10 7				
	MealthyControl			8				
				6 -			1.1.1.1	
	ان ال ال			4-1			սել, հեմ է հե	
	i at third k a differenti h	ora n in the the the the t	hallata di shiftini	an a	tit bit di lattar	VER DE TREBENE	lill i shidhini i lil	hili ali bi ali 1