



Ingebiol

Diego Darío Guerrero Fernández



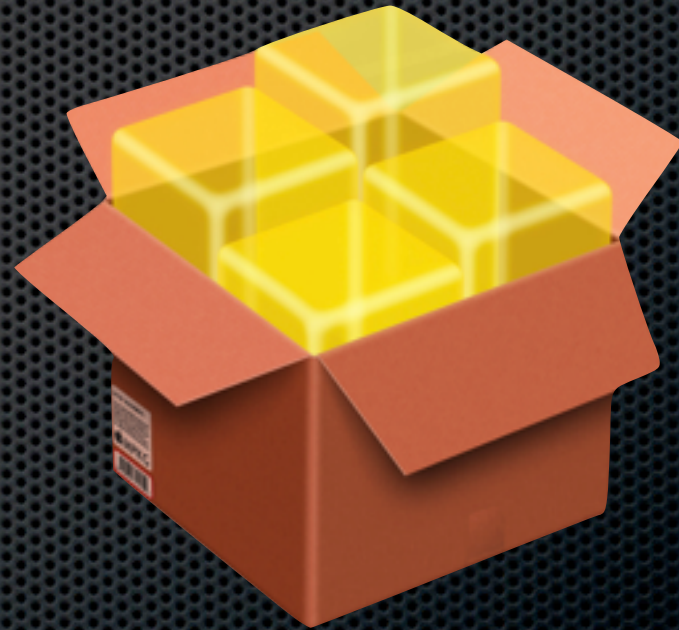
PLATAFORMA
ANDALUZA DE
BIOINFORMÁTICA



UNIVERSIDAD
DE MÁLAGA

Contenido

- ✧ ¿Qué es Ingebiol?
- ✧ Definición de un flujo de ejemplo: **GEN**ote
- ✧ Preguntas



¿Qué es Ingebiol?

¿Qué es Ingebiol?

- **Generador parametrizado de interfaces gráficos:**

- Generador parametrizado de interfaces gráficos.
- Soporte para el envío a sistemas de colas para supercomputación.
- Configuración basada en ficheros de texto.
- Capacidad para gestionar flujos de trabajo.
- Posibilidad de usar cualquier programa en línea de comandos sea cual sea su lenguaje.

¿Cómo funciona?

- ✦ **Intervienen dos tipos de usuarios:**



Administradores

- Definir:
 - Parámetros de entrada
 - Los flujos de trabajo
 - Salida de los flujos



Usuarios finales

- Envían trabajos
- Reciben resultados

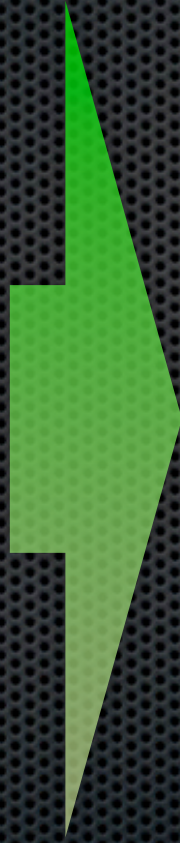
El administrador

- Define los parámetros de entrada:

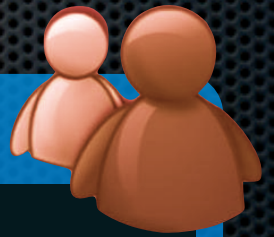
Secuencias

Identidad
mínima

Otros
parámetros



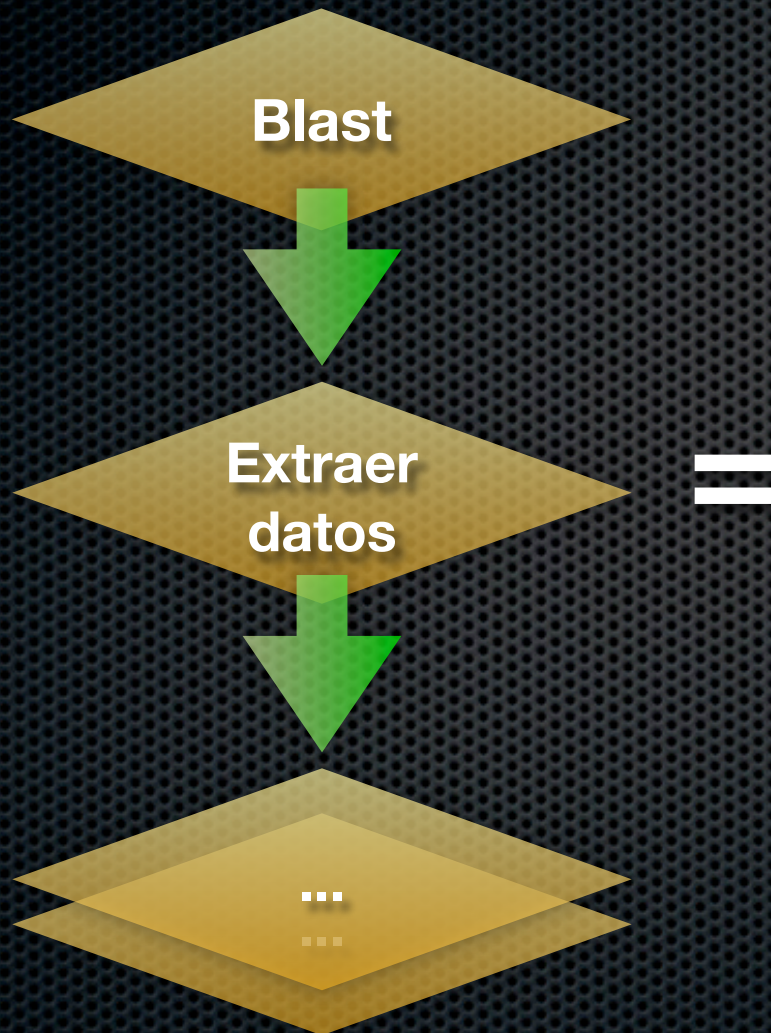
Administrador



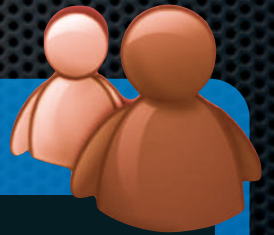
```
{
  input_type: 'file',
  id: 'input_file',
  save_as: 'input_file.fasta',
  title: 'Sequences (fasta)',
  default_value: '',
  required: true,
  tooltip: 'Select a sequence file in fasta format'
},
{
  input_type: 'integer',
  id: 'min_ident',
  command_switch: '%s',
  title: 'Minimum identity (%)',
  required: true,
  tooltip: 'The minimum identity required to..., in percentage.',
  default_value: '45',
  validation_lower_limit: '45',
  validation_upper_limit: '100'
}
```


El administrador

- ✦ Define el flujo de trabajo:



Administrador



```
command_list:[
  {
    required_files:[],
    command:'wublast input_file.fasta -o wublast_results.txt'
  },
  {
    required_files:['input_file.fasta','wublast_results.txt'],
    command:'genote.rb'
  }
]
```

Posibilidad de usar cualquier programa en línea de comandos sea cual sea su lenguaje.

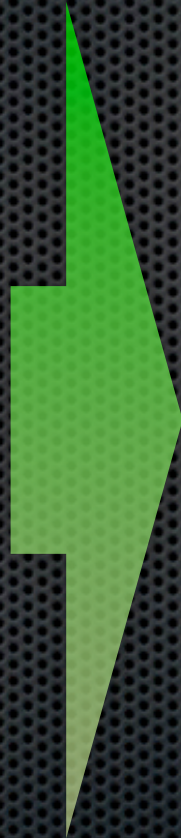
El administrador

- Define los elementos de salida:

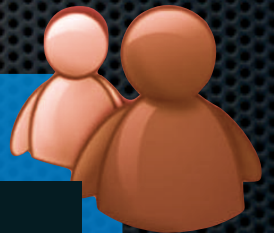
Resultados

Imágenes

Otros
ficheros



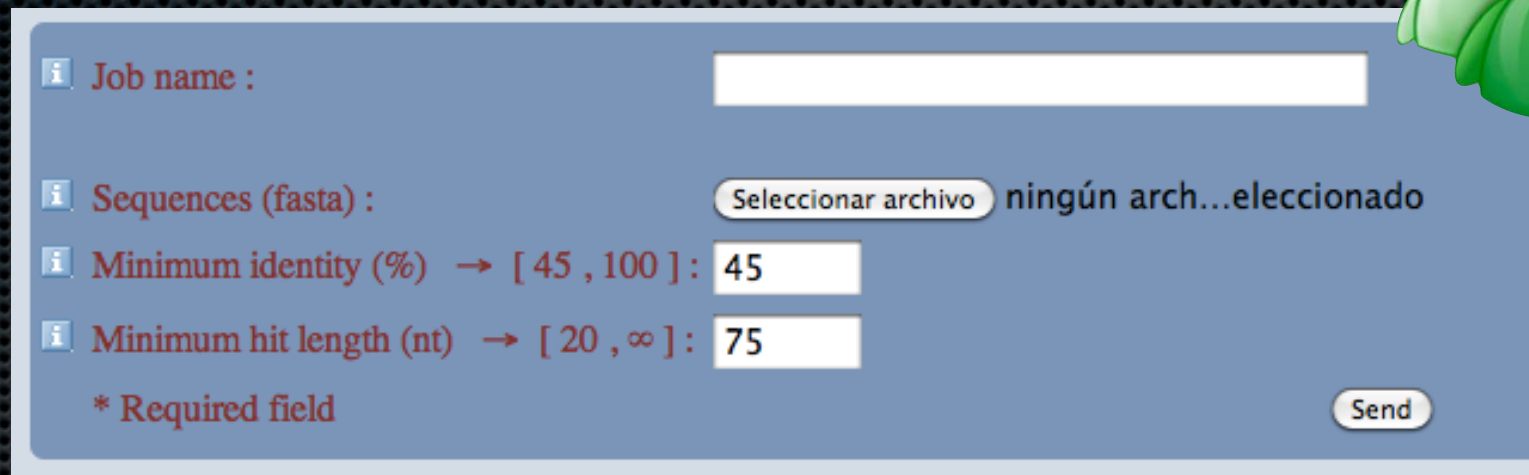
Administrador



```
{
  type: 'file',
  file: 'putative_consecutive_query.txt',
  title: 'Putative consecutive query',
  tooltip: ''
},
{
  id: 'image1',
  type: 'image',
  file_name: 'gff_regions.pdf',
  content_type: 'image/pdf',
  title: 'GFF view',
  options: {height: '200', width: '250'},
  linked: true,
  tooltip: ''
}
```


El usuario final

- ✦ **Rellena una página web con los valores de los parámetros**



The screenshot shows a web form with a light blue background. It contains several input fields and a button. To the right of the form, there are two green 3D-style icons of people. The form fields are as follows:

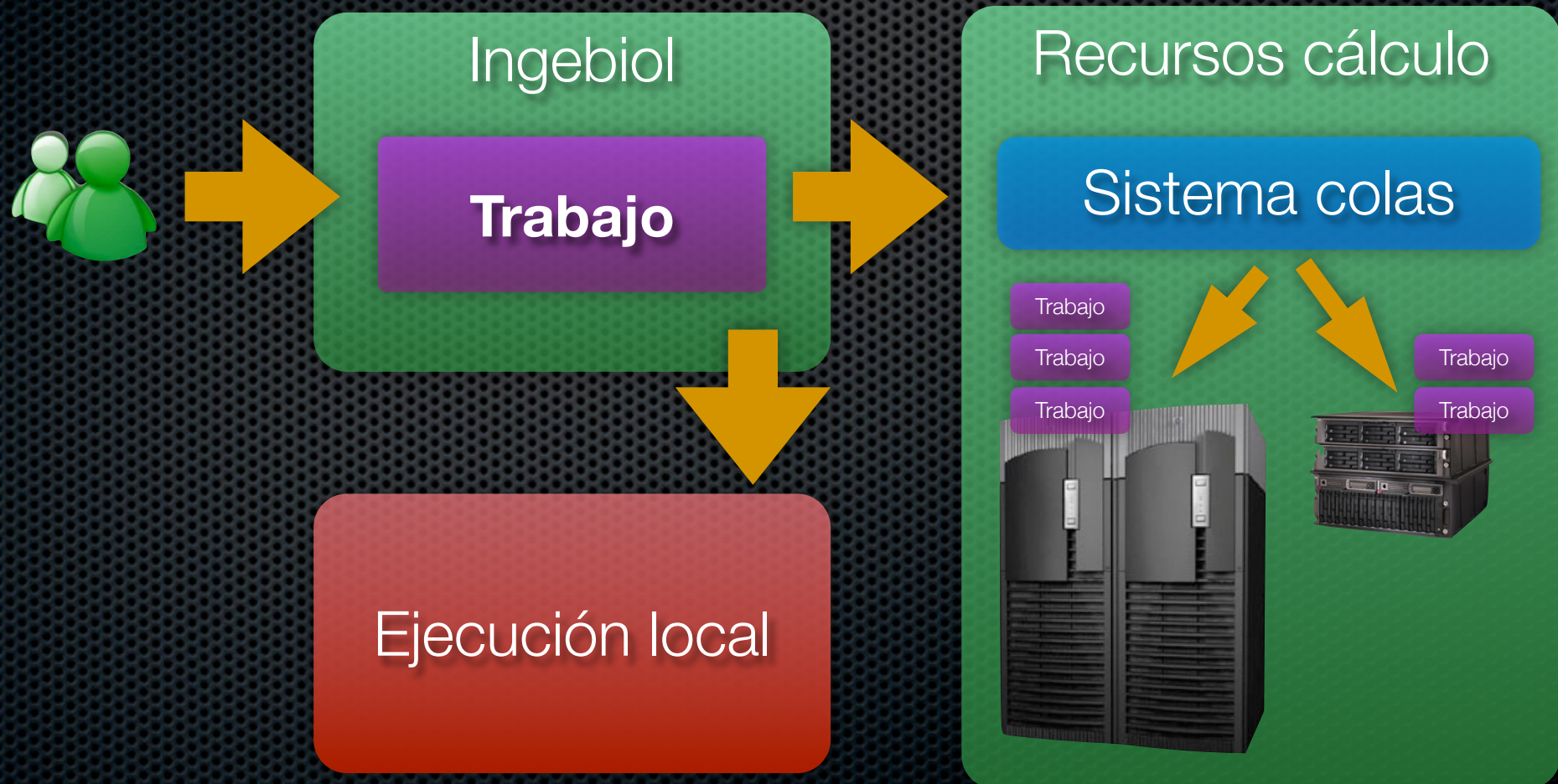
- Job name :** A text input field.
- Sequences (fasta) :** A button labeled "Seleccionar archivo" followed by the text "ningún arch...eleccionado".
- Minimum identity (%) → [45 , 100] :** A text input field containing the value "45".
- Minimum hit length (nt) → [20 , ∞] :** A text input field containing the value "75".

At the bottom left of the form, there is a note: "* Required field". At the bottom right, there is a button labeled "Send".

- ✦ **Realiza el envío del trabajo**

El usuario final

- La ejecución es transparente para el usuario y puede ir a sistemas de colas para super-computación:





El usuario final

- ✦ **Revisa los trabajos que había enviado**

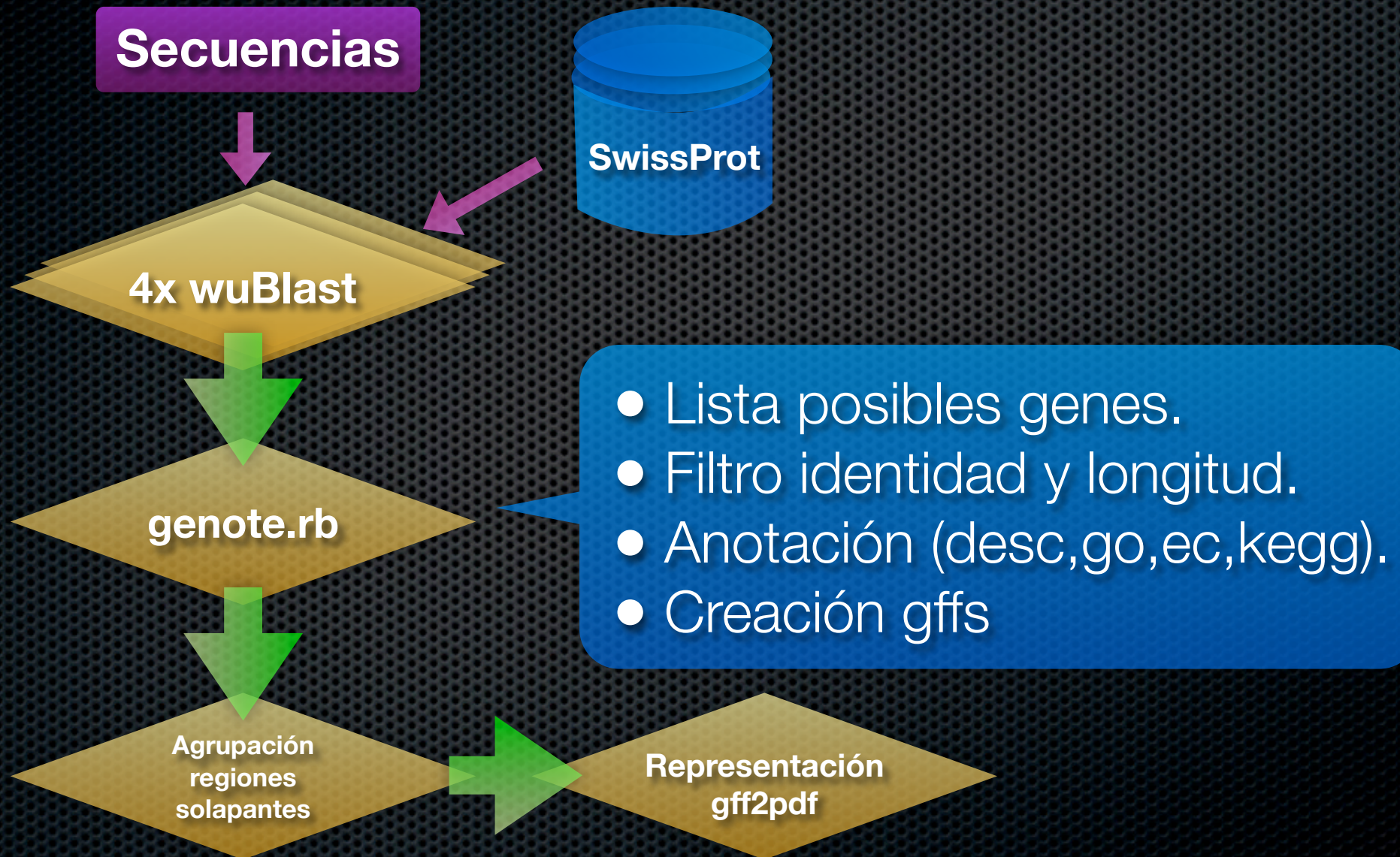
JOB LIST					
Job name	Job size	File name	Status	Params	Refresh
job1	308K		DONE		
multi	308K		DONE		

- ✦ **Visualiza el trabajo que desea y descarga los resultados:**

RESULTS	
	Regions (GFF) (1.3KB)
	Regions (PDF) (4.6KB)
	Annotated results (45.9KB)
	Complete (GFF) (17.8KB)

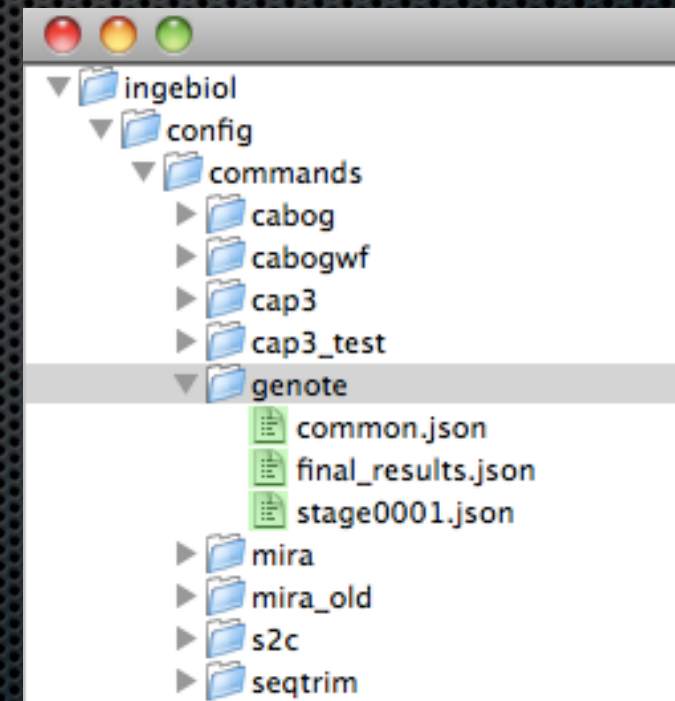
Ejemplo de flujo

GENote



GENote

- ✧ Ingebiol puede gestionar multiples “flujos”.
- ✧ Cada “flujo” posee su propia configuración.
- ✧ Ficheros formato JSON (legibles):
 - ✧ 1 common
 - ✧ 1 final_results
 - ✧ N stages



GENote - Entrada

- **Secuencias (file)**
- **Identidad min (int)**
- **Longitud min (int)**

```
input_params: [  
    param1,  
    param2,  
    param3  
]
```

```
##### PARAM1 #####  
{  
    input_type: 'file',  
    id: 'input_file',  
    save_as: 'input_file.fasta',  
    title: 'Sequences (fasta)',  
    default_value: '',  
    required: true,  
    tooltip: 'Select a sequence file in fasta format'  
}  
  
##### PARAM2 #####  
{  
    input_type: 'integer',  
    id: 'min_ident',  
    command_switch: '%s',  
    title: 'Minimum identity (%)',  
    required: true,  
    tooltip: 'The minimum identity required to...',  
    default_value: '45',  
    validation_lower_limit: '45',  
    validation_upper_limit: '100'  
}
```


GENote - Flujo

- ❖ **wuBlast**
- ❖ **genote**
- ❖ **gff2pdf**

```
command_list: [  
    command1,  
    command2,  
    command3  
]
```

```
##### COMMAND1 #####  
{  
    required_files:['sp_db','input_file.fasta'],  
    command:'blastx sp_db input_file.fasta -  
sort_by_totalscore -hspmax=0 -E=1e-10 -mformat=3 -  
cpus 4 -o wublast_results.txt'  
}  
  
##### COMMAND2 #####  
{  
    required_files:['input_file.fasta',  
'wublast_results.txt', 'sp_anotaciones.txt'],  
    command:'genote.rb $min_ident $min_hit_length'  
}  
  
##### COMMAND3 #####  
{  
    required_files:['gff_regions.gff'],  
    command:'gff2ps -a -c -v -C gff2psrc  
gff_regions.gff > gff_regions.ps  
2>gff2ps_errors.txt'  
}
```


GENote - Resultados

- **gff_regions (file)**
- **pdf (file)**
- **Result_annotated (file)**

```
params: [  
    result1,  
    result2,  
    result3  
]
```

```
##### RESULT1 #####  
{  
  type: 'file',  
  file: 'gff_regions.gff',  
  title: 'Regions (GFF)',  
  tooltip: ''  
}
```

```
##### RESULT2 #####  
{  
  type: 'file',  
  file: 'gff_regions.pdf',  
  title: 'Regions (PDF)',  
  tooltip: ''  
}
```

```
##### RESULT3 #####  
{  
  type: 'file',  
  file: 'Result_annotated.txt',  
  title: 'Annotated results',  
  tooltip: ''  
}
```


FIN