

# Trabajo Tutelado

Curso 2021-2022

## Computación de Altas Prestaciones en Bioinformática Master Universitario en Bioinformática para Ciencias de la Salud

El objetivo de este trabajo tutelado es que el alumno se familiarice con la forma de trabajar en entornos paralelos compartidos por varios usuarios y descubra de primera mano los beneficios que puede obtener de estos sistemas en el ámbito de la bioinformática. Como objetivos secundarios se pretende que alumno adquiera competencias en el desarrollo, escritura y exposición de trabajos de investigación relacionados con la computación de altas prestaciones. El trabajo tutelado será **individual** y constará de las siguientes fases:

1. **Selección de la herramienta/aplicación objeto de estudio.** El alumno deberá realizar una búsqueda online para encontrar herramientas bioinformáticas que hagan uso de la capacidad de cómputo de sistemas paralelos para acelerar sus cálculos. Las herramientas deben poder ejecutarse en el supercomputador Finis Terrae II del CESGA. Por tanto, deben estar disponibles públicamente para su descarga con una licencia de código abierto. Se admiten tanto paralelizaciones en sistemas multinúcleo como aceleradores hardware (en ambos casos con uno o varios nodos). En cuanto al ámbito de la bioinformática, no se impone ninguna restricción, pudiéndose centrar en el área de genética, genómica, análisis de secuencias, etc. Para esta tarea puede ser útil emplear Google Scholar (<http://scholar.google.com>) y/o bucear en artículos de revistas de investigación. Algunos ejemplos son:

- Bioinformatics (<http://bioinformatics.oxfordjournals.org/>)
- BMC Bioinformatics (<https://bmcbioinformatics.biomedcentral.com/>)
- BMC Genomics (<https://bmcbgenomics.biomedcentral.com/>)
- BMC Genetics (<https://bmcbgenet.biomedcentral.com/>)

El alumno dispondrá de tres semanas para esta tarea de investigación inicial, debiendo presentar al profesor **tres sugerencias** de aplicaciones bioinformáticas paralelas **como muy tarde** al inicio de la **clase de prácticas del día 08/03**. Se puede (y se recomienda) presentar con anterioridad bien en tutorías (pidiendo cita) o en clases previas. En el momento de la presentación el alumno debe entender el objetivo de cada una de las propuestas así como el hardware que requieren. El profesor ayudará al alumno a decantarse por una de las mencionadas herramientas.

2. **Instalación, ejecución y obtención de resultados.** Una vez decidida la herramienta sobre la que se centrará el trabajo, el alumno debe proceder a su descarga, instalación y ejecución en su cuenta del Finis Terrae II. Cabe destacar que no todas las herramientas son fáciles de instalar y/o ejecutar, así que el alumno puede necesitar la ayuda del profesor o incluso puede ser

necesario requerir información adicional a los autores de la herramienta y/o el servicio de soporte del CESGA. Una vez el alumno se haya familiarizado con la herramienta deberá trazar un plan de pruebas que permita determinar la escalabilidad de la solución paralela. Este plan de pruebas debe incluir datos como el tipo y tamaño de los conjuntos de entrada a analizar, el número de núcleos o aceleradores hardware a emplear, etc. Este plan de pruebas debe ser presentado al profesor como muy tarde durante la **clase de prácticas del día 22/03**, y siempre antes de empezar las pruebas en sí. De esta forma se podrán detectar errores en la planificación antes de que eso lleve a una pérdida de tiempo y recursos innecesarios. Se recomienda entregar este plan de pruebas en una fecha anterior a la obligatoria, bien usando las clases de prácticas o bien mediante comunicación por correo electrónico con el profesor.

3. **Escritura de un trabajo de investigación con los resultados.** El alumno deberá incluir una breve descripción de la herramienta así como los resultados obtenidos con ella en un trabajo de investigación. Es muy importante remarcar que, salvo causas de fuerza mayor previamente autorizadas por el profesor, los resultados experimentales presentados en este trabajo deben ser los obtenidos en el Finis Terrae II. Nunca se deben presentar como propios los resultados experimentales que puedan aparecer en el artículo original donde se presentó la herramienta (el obtenido en la fase 1 de este trabajo). En cuanto al formato para el artículo, se deben tomar todas las consideraciones precisas como si se fuese a presentar una *application note* al journal Bioinformatics (<http://bioinformatics.oxfordjournals.org/>). La información acerca del formato, así como plantillas de *LaTeX* para ayudar a la escritura, pueden obtenerse en: [http://www.oxfordjournals.org/our\\_journals/bioinformatics/for\\_authors/general.htm](http://www.oxfordjournals.org/our_journals/bioinformatics/for_authors/general.htm). **AVISO! Es muy importante seguir el formato de la revista. Se penalizarán errores como no ajustarse al tamaño máximo, usar una única columna, posición incorrecta de las captions de tablas, formato de bibliografía incorrecto, etc.** La mejor solución para comprobar que el formato es correcto es comparar tu artículo con uno ya publicado en la revista. Se recomienda consultar cualquier duda respecto al formato con el profesor de la asignatura. La fecha límite para la **entrega del artículo es el viernes 08/04 a las 23:59 horas**. La entrega se hará a través del Campus Virtual y solo será necesario entregar el archivo pdf final.
4. **Presentación del trabajo ante el profesor y el resto de la clase.** El alumno deberá desarrollar y preparar una presentación para dar a conocer su trabajo en horario de clase. Cada alumno dispondrá de **15 minutos para su presentación y 5 minutos para preguntas**, tanto del profesor como del resto de sus compañeros. Es importante adecuar la presentación al tiempo disponible, sin quedarse corto ni pasarse del tiempo. Las presentaciones se harán en horario de clase **entre el 19/04 y el 12/05** (orden aleatorio del que se informará más adelante). La asistencia a las presentaciones del resto de compañeros es obligatoria.

Se recuerda que la elaboración en inglés tanto del trabajo escrito como de la presentación puede desembocar en un punto en la nota del alumno.