



e.Rrados

ENTREVISTA A UNA CIENTÍFICA – R PARA LA VIDA

**ENTREVISTA REALIZADA A LA**

**DR. DOÑA DANIELA LEDEZMA TEJEIDA**

**ETH ZÜRICH**

**PROFESSUR FÜR SYSTEMBIOLOGIE**

**-Diciembre 2020-**

**1) ¿Cuál es el papel que ocupa R en la investigación científica hoy en día?**

Desde mi experiencia en el área de biología computacional R facilita el manejo, procesamiento y análisis de datos masivos. En especial en el análisis estadístico. Por su carácter open-source permite la colaboración indirecta de miles de investigadores de disciplinas distintas, de forma que un experto estadístico puede generar un paquete que ayude a un experto en biología a analizar sus datos de forma sencilla.

**2) ¿Qué ventajas presenta frente a otros sistemas como Matlab? ¿Qué desventajas?**

La ventaja más importante contra MatLab es que R es un software de libre distribución, lo que permite que cualquier investigador pueda utilizarlo, sin importar el presupuesto que tenga para su investigación. De la mano va que existan muchos cursos online o gratuitos (como los proyectos Software Carpentry y DataCarpentry) para que cualquiera pueda aprender a usarlo. Asimismo, cualquier desarrollador puede generar y publicar paquetes para temas muy específicos, de los que otros usuarios pueden beneficiarse.

La mayor desventaja sería que los estándares para la documentación de paquetes son muy laxos, de forma que algunos paquetes se tornan difíciles de usar porque sus manuales no son claros. En este tema Matlab sin duda tiene mucha ventaja.

**3) ¿Podría contarnos un poco cuál es el objeto de su investigación científica? ¿Qué papel desempeña R en ella?**

Mi investigación está enfocada la relación entre la regulación transcripcional y el metabolismo de bacterias, tanto en términos funcionales como mecánicos. En mi investigación genero cantidades masivas de datos de expresión genética y metabólica. R me permite procesarlos y analizarlos.

**4) ¿Cree usted que la formación universitaria en los planes actuales ofrece una base sólida en conocimientos de Bioinformática y manejo de programas como R?**



e.Rrados

## ENTREVISTA A UNA CIENTÍFICA – R PARA LA VIDA

No cuento con suficiente información para dar una respuesta contundente dado que mi programa de formación universitaria estaba específicamente enfocado a bioinformática. En mi limitada experiencia, creo que los programas de biología a nivel universidad poco a poco han incluido más temas de bioinformática tales como principios de programación. No creo que debería ser obligatorio incluir cursos de bioinformática en cualquier licenciatura de biología. Sospecho que hay áreas de especialización en biología que no requieren el manejo de datos masivos. Creo que las maestrías serían un gran lugar para empezar a ofrecer opciones mucho más enfocadas a bioinformática.

### **5) ¿Qué ejemplos de la vida cotidiana o de la investigación científica básica utiliza R?**

Cualquier evento de la vida cotidiana o de investigación que requiera usar Microsoft Excel, o una calculadora, puede ser resuelto en R. El ejemplo más cotidiano que se me ocurre es que durante la pandemia algunos colegas bioinformáticos generamos código en R para poder realizar juegos por zoom que requerían el uso de datos o cualquier otra asignación al azar.

### **6) ¿Qué considera imprescindible saber a la hora de manejar R? ¿Cuál es la mejor manera de aprender?**

Idealmente, tener nociones básicas de programación. No en algún lenguaje en específico, simplemente ser capaz de generar pseudocódigo para tener claro cuales son los requisitos para resolver el problema al que uno se enfrenta. Todo lo demás es aprender comandos que se pueden googlear fácilmente.

La mejor manera de aprender es con la práctica. Buscando problemas básicos de programación y resolverlos haciendo diagramas de flujo en papel.

### **7) ¿Considera usted importante saber desenvolverse en R para aquellas personas que se estén formando en Biotecnología?**

Creo que R es una herramienta, tan importante como muchas otras. Creo más importante que el alumno tenga una idea de qué áreas de biotecnología son de su interés y si esas áreas utilizan datos masivos. Si es el caso, entonces considero importante aprender principios básicos de programación. De esta forma, si de pronto se necesita usar Python en lugar de R, o el alumno llega a un laboratorio donde sólo existe Matlab, la transición será mucho más fácil.



e.Rrados

ENTREVISTA A UNA CIENTÍFICA – R PARA LA VIDA

**8) ¿En qué estado se encuentra la Bioinformática a día de hoy? ¿Qué futuro le ve?**

Creo que la bioinformática está avanzando aceleradamente. Cada vez somos capaces de generar más datos y creo que la frontera actualmente está en integrar datos de distintas tecnologías -ómicas utilizando machine learning.

En la bioinformática más clásica existe la idea de que la estadística es la disciplina que nos permite encontrar sentido a los datos, de ahí la relevancia de R. En mi opinión, la estadística permite encontrar los datos relevantes, pero se requiere de personas que sean capaces de tomar los resultados estadísticamente significativos y devolverlos al contexto biológico, ese es el verdadero reto. Creo que en el futuro se valorarán más las herramientas y aplicaciones que permiten extraer conocimiento de los datos, y no aquellas que sólo los transformen. En un ejemplo más aterrizado, los análisis de expresión diferencial de genes (RNA-seq, microarreglos, etc.) producen listas de genes con cambios de expresión significativos que usualmente terminan convertidas en listas de Gene Ontologies. Se requieren herramientas que nos digan cómo se relacionan las Gene Ontologies entre ellas y qué significan esas relaciones. De otra forma, seguiremos leyendo artículos que comienzan con millones de datos y terminan validando 1 ejemplo interesante.