



e.Rados

ENTREVISTA A UN CIENTÍFICO – R PARA LA VIDA
ENTREVISTA REALIZADA AL SEÑOR

DON RODRIGO ESPINOZA SILVA

ANALISTA –

**DIAGNÓSTICO MOLECULAR EN DÍAZ GILL
MEDICINA LABORATORIAL**

-Diciembre 2020-

1. ¿Cómo de importante resultan los conocimientos de programación e informática en la investigación biotecnológica o científica en general?

Voy a citar un artículo que me llegó mucho cuando lo leí: "All biology is computational biology" del PhD Florian Markowetz. Se puede encontrar de manera gratuita en la revista PLOS Biology. Las capacidades informáticas de un científico en el área de la biología y afines, ya ha dejado de ser un "plus", y se ha vuelto una competencia básica. Dependiendo del tipo de investigación que uno haga, se vuelve más necesario o no el uso de herramientas bioinformáticas, pero todas ellas necesitan en algún punto, analizar grandes cantidades de datos, que sólomente podrán ser factibles en un procesador. Para ello existe una gran variedad de lenguajes, pero los más recomendados para aprender a programar son Python y R. Ambos son muy utilizados en el ámbito de las ciencias biológicas, por su facilidad de aprendizaje, su modo intuitivo de operar, y porque tiene una comunidad muy activa que está actualizando los paquetes y siempre disponible ante cualquier problema que vaya surgiendo en el uso.

Actualmente, querer investigar en cualquier ciencia biológica sin saber usar los softwares y las herramientas informáticas, es como correr a caballo en una carrera contra automóviles. Dicho esto, muchas herramientas informáticas tienen una página en internet, donde uno puede acceder y utilizarlas sin conocimientos de programación, aunque son mucho más limitadas que su uso desde la línea de comandos o un lenguaje como R o Python.

2. ¿Le daría la misma o incluso más importancia con respecto a otras asignaturas típicamente impartidas en la educación secundaria y superior?



e.Rrados

ENTREVISTA A UN CIENTÍFICO – R PARA LA VIDA

Yo he cursado materias que todavía no entiendo con qué finalidad se agregaron a la malla curricular de un biotecnólogo, mientras tantas otras, como programación en Python, en R, manejo de bases de datos, por un lado, o matemáticas (álgebra lineal, principalmente) y estadísticas aplicadas a las ciencias genómicas, por otro lado, son fundamentales para la investigación biológica y no se dan de forma curricular, al menos en mi país.

3. ¿Qué usos de R en investigación científica ha visto o vivido en a lo largo de su experiencia laboral? ¿Qué usos profesionales de bioinformática (y de R, si es el caso) está usted llevando a cabo actualmente?

Lo que me tocó usar R en investigación, es lo que pude compartir en el curso que tuvimos en el marco de los 10 años de la licenciatura en Biotecnología de mi facultad: análisis de expresión diferencial con datos de RNA-seq y, ahora mismo, estoy aprendiendo a usar paquetes de R que realizan estudios eQTL (en inglés, Expression Quantitative Trait Loci), donde se correlaciona variaciones puntuales (SNPs) con expresión diferencial en genes, y quiero aplicar a la sintomatología de COVID, para identificar mutaciones puntuales que puedan influir en la severidad de síntomas (QTL) y/o en la expresión de genes (eQTL), relacionados a esta enfermedad. Un análisis como eQTL está correlacionando cientos de miles de SNPs, contra decenas de miles de genes en cientos de pacientes, algo que sería imposible realizar sin un ordenador.

4. ¿Qué desearía haber sabido antes de terminar la carrera y entrar en el mundo de la investigación científica, y por qué?

Me hubiese gustado sumergirme mucho más en las matemáticas y estadísticas que respaldan los algoritmos computacionales para analizar datos biológicos. De esta manera se entiende mucho mejor qué se está realizando con X herramienta, y cómo modificar parámetros para adecuar la herramienta a lo que uno quiera analizar. Sin un fondo matemático sólido, estas herramientas se vuelven muy difíciles de entender, y el análisis de los datos puede terminar siendo erróneo por no entender, por ejemplo, las asunciones que se toman para aplicar un cierto modelo matemático con un cierto tipo de datos.

Además, la parte administrativa de la gestión de proyectos de investigación también es algo que uno recién al entrar al mundo laboral o profesional empieza a tener contacto, y es algo que hubiese querido aprender de forma más práctica durante mis estudios.



e.Rrados

ENTREVISTA A UN CIENTÍFICO – R PARA LA VIDA

5. ¿Cuáles cree que son los aspectos a mejorar del plan de estudios en el cual se formó para una mejor preparación en bioinformática?

Como ya dije previamente, las matemáticas y estadísticas son materias que se deben reforzar en un plan de estudios de un biólogo o biotecnólogo. En mi caso, bioinformática era una materia electiva, lo cual ya no me parece que debe ser el caso. Un biotecnólogo actualmente necesita tener algún conocimiento informático para ser competitivo en el mundo laboral, si quiere dedicarse a la investigación. Creo que también, debemos salir de la facultad con al menos un lenguaje de programación dominado.

6. ¿Cómo ve el futuro de la bioinformática en el medio plazo?

La bioinformática se va a seguir expandiendo, y se va a ir instalando cada vez más para ocupar una rama central en la biología. Los expertos en bioinformática serán cada vez más demandados en el mercado ya que podrán aplicar sus conocimientos a cualquier ciencia biológica, dentro y fuera de la academia. (es por esto, además de que me apasiona la bioinformática, que quiero seguir especializándome en ésta).

7. ¿Cuáles son las dificultades que se ha encontrado en el desarrollo de su actividad profesional en relación a la bioinformática?

Como es una rama de la ciencia muy nueva, hay muy pocos investigadores en Paraguay que manejen la bioinformática desde sus fundamentos. Sí hay muchos investigadores que saben usar las herramientas pero a nivel muy empírico. Entonces cuando uno intenta formarse en este área, no encuentra muy fácilmente a un tutor apropiado, y finalmente termina aprendiendo empíricamente también.

8. ¿Qué potencialidades o ventajas presenta R en el campo de la Biotecnología frente a otros lenguajes de programación? ¿Qué inconvenientes?

El lenguaje R tiene una interfaz gráfica que le hace mucho más fácil de usar, llamada RStudio. Esto permite que muchos científicos que no precisamente dominen la informática, puedan usar las herramientas que ofrece R. Además, tiene dos ventajas que se me ocurren en este momento: los llamados "paquetes" en R, son herramientas que uno va descargando de acuerdo a el uso que necesite en el momento, lo cual da al



e.Rrados

ENTREVISTA A UN CIENTÍFICO – R PARA LA VIDA

usuario una paleta de miles de funciones ya preparadas por científicos del área, y no tiene que esforzarse en programar cada función que vaya necesitando en su investigación. Por otra parte, los gráficos que se pueden producir en R son muy estéticos y listos para ser publicados, y nuevamente, no son muy difíciles de aprender a generar.

9. ¿Cuál es su percepción respecto a la conciencia de la población general y de la comunidad científica de la importancia creciente de la Bioinformática?

Desde mi perspectiva, los estudiantes están cada vez más interesados en la bioinformática, pero la ven como una materia muy lejana o difícil, porque combina mucha matemática e informática, lo cual puede ser muy abrumante en un principio. Entonces no muchos estudiantes se animan a incursionar en esta rama tan apasionante y necesaria.

10. Finalmente, ¿qué otros aspectos considera de relevancia comentar en relación al lenguaje de programación R aplicado a la Biotecnología?

Todas las aplicaciones que estuve comentando previamente, se centran en el uso de R para análisis de "Big Data" biológico. Pero este lenguaje no se limita a su uso biológico. Es una herramienta estadística aplicable a cualquier área, y, como biotecnólogos, nosotros debemos manejar también otras partes del proceso de producción de un producto, por ejemplo. Los análisis financieros que envuelven a un proyecto, sea de investigación o industrial, también pueden aplicarse en R.