News Letter

What is Bioinformatics?

Andres Barriga Gonzalez*

Department of Chemical Sciences and Pharmacy, University of Chile, Chile

*Corresponding author

Andres Barriga Gonzalez, Department of Chemical Sciences and Pharmacy, Mass Spectrometry Unit, University of Chile, Chile, Tel: 569229782991; Email: anbarr@ciq.uchile.cl

NEWS LETTER

It is the application of computational tools for the resolution of a biological problem based on different analytical measurements, they can be online tools that require the upload of data or they can be downloadable offline tools. There are no universal tools but they focus on a specific biological area (some omics). In addition, the tools can be classified as commercial or non-profit.

The Mass Spectrometry Unit of the Faculty of Chemical Sciences and Pharmacy of the University of Chile is a service laboratory for academic research where our work is essentially focused on the bottom-up identification of proteins (proteomics service), that is, protein identification from tryptic peptides; in the characterization/sequencing of peptides (peptidomics service) and in the identification of various natural products (metabolomics service). Identification proteomic analysis is based on the use of Mascot online tool (http://www.matrixscience.com/search_form_select.html) which is generally satisfactory in the identification of proteins that have similarity or homology with those protein sequences contained in databases such as NCBI or Swiss Prot, however, you cannot add protein sequences or genomic sequences of species not included in the databases or do not have de novo sequencing capability, options that are found in the commercial version. Alternative to Mascot, is PEAKS Studio (http://www.bioinfor.com) that presents the option of de novo sequencing and addition of protein and/or genomic sequence but it is a completely commercial tool. To our knowledge to date non-commercial tools similar to Mascot or PEAKS Studio have not been developed.

Within the peptidomic analysis we distinguish those natural ribosomal peptides as well as those that originate from the natural proteolysis of proteins (both linear peptides) or those non-ribosomal peptides that can be linear, cyclic, branched, mixed, etc. In the case of ribosomal peptides there are few de novo sequencing tools which allow not defining a protease with the exception of PepNovo (http://proteomics.ucsd.edu/Software/PepNovo/). For non-ribosomal peptides there is a very good tool named CycloBranch (http://ms.biomed.cas.cz/cyclobranch/) which allows analyzing a great variety of peptide forms either by using a database or by allowing de novo sequencing.

For the identification analysis of secondary metabolites or natural products there is a wide range of tools both online as METLIN (https://metlin.scripps.edu/), MetFrag (https://msbi.ipb-halle.de/MetFrag/), MassBank (http://www.massbank.jp), ReSpect (http://spectra.psc.riken.jp/) among others as desktop tools such as Sirius (https://bio.informatik.uni-jena.de/software/sirius/), MZmine 2 (http://mzmine.github.io/) and CFM-ID (https://sourceforge.net/projects/cfm-id/). There are specific tools for some families of compounds such as those for the analysis of lipids.

Independent of the analysis to be carried out the bioinformatics tools tend to present higher quality results for those data obtained from high or ultra resolution instruments.

Many of these tools can read the original raw format produced by the instruments, however, others need formats such as mzXML, mgf (generic Mascot format) or similar, several brands of instruments provide format converters, but others do not.

Where to obtain bioinformatics tools?

Among the databases we usually consult in search of bioinformatics tools are SourceForge (https://sourceforge.net/), GitHub (https://github.com/) and OMICtools (https://omictools.com/).

What would be my ideal bioinformatics tool?

It does not necessarily mean that it will analyze all kinds of samples but as a simple user without knowledge as a developer or programmer you would expect a simple tool because some request many parameters of analysis most of them unknown to a beginner user and even to an advanced user and that would provide more information simultaneously as in the case of peptides where it is unknown a priori if these are linear, cyclic, branched, etc., which translates into several analyzes must be performed. The bioinformatic tools used for the identification of metabolites are mainly based on the comparison of fragmentation data in high resolution (TOF analyzers) or ultra-resolution (Orbitrap and FT-ICR) with servers that generate in silico fragmentations based on structures contained in different bases of data or in some cases the tools compare the experimental fragmentation data with fragmentation libraries obtained at different fragmentation voltages [1]. An interesting option would be the ability of de novo interpretation of fragmentation data considering that there is abundant literature on the fragmentation of several families of compounds, such as glucosinolates [2], flavonoids [3], carotenoids [4], alkaloids [5], pentacyclic triterpene acids [6,7] as well as some of their modifications as glycosylations [1]. This

SciMedCentral

approach could be applied independently of the resolution of the data.

Undoubtedly, bioinformatics tools have facilitated the interpretation of experimental data as well as being able to analyze a large volume of these.

REFERENCES

1. Cao J, Yin C, Qin Y, Cheng Z, Chen D. Approach to the study of flavone di-C-glycosides by high performance liquid chromatography-tandem ion trap mass spectrometry and its application to characterization of flavonoid composition in *Viola yedoensis*. J Mass Spectrom. 2014; 49: 1010-1024.

2. Cataldi TR, Lelario F, Orlando D, Bufo SA. Collision-induced dissociation of the A+2 isotope ion facilitates glucosinolates structure elucidation by electrospray ionization-tandem mass spectrometry with a linear quadrupole ion trap. Anal Chem. 2010; 82: 5686-5696.

3. Fabre N, Rustan I, de Hoffmann E, Quetin-Leclercq J. Determination of flavone, flavonol, and flavanone aglycones by negative ion liquid chromatography electrospray ion trap mass spectrometry. J Am Soc Mass Spectrom. 2001: 12: 707-715.

4. Rivera SM, Christou P, Canela-Garayoa R. Identification of carotenoids using mass spectrometry. Mass Spectrom. Rev. 2014; 33: 353-372.

5. Shim H, Lee JY, Kim B, Hong J. General fragmentations of alkaloids in electrospray ionization tandem mass spectrometry. Mass Spectrom Lett. 2013; 4: 79-82.

6. Syed Ghulam Musharraf, Nayab Kanwal, Qamar ul Arfeen. Stress degradation studies and stability-indicating TLC-densitometric method of glycyrrhetic acid. Chem Cent J. 2013; 7: 9.

7. Xia B, Bai L, Li X, Xiong J, Xu P, Xue M. Structural analysis of metabolites of asiatic acid and its analogue madecassic acid in zebrafish using LC/IT-MSn. Molecules. 2015; 20: 3001-3019.