Roberto Vera Alvarez, PhD

E-mail: Roberto.Vera@icgeb.org, r78v10a07@gmail.com
Phone Number: +39-389-655-5220 (Italian mobile).

Personal Data

Cuban citizen, male, born in 1978.

Education

2013 - 2014	PhD in Bioinformatics, Pázmány Péter Catholic University, Budapest, Hungary.
2001 - 2003	Bioinformatics Specialist, CIGB, Havana, Cuba.
1997 - 2003	Degree in Physics, University of Havana, Cuba.

Other Courses

2009	Bioinformatics: Computer Methods in Molecular Biology, ICGEB, Trieste, Italy.
2007	Latest Advances in Drug Discovery Design & Planning Methods, Oxford University, UK.
2004	Molecular Design Advance Seminar and Bioinformatics, University of Havana, Cuba.

Career History

2015 -	IT Consultant, Biosafety Group, ICGEB, Trieste, Italy.
2013 - 2014	Research Fellow, LNCIB, Trieste, Italy.
2012 - 2013	Research Fellow, Protein Structure and Bioinformatics Group, ICGEB, Trieste, Italy.
2010 - 2012	Assistant Professor, Physic Department, ISPJAE, Havana, Cuba.
2003 - 2010	Bioinformatics Specialist, CIGB, Havana, Cuba.

Teaching experiences

2014	Bioinformatics on the Google Cloud Platform, ICGEB.
2006 - 2012	Teach Classical Physics at Physic Department, ISPJAE.
2004	Parallel Programming for Bioinformatics. Bioinformatics Specialist Course, second edition, CIGB.

Languages

- English (fluent), Italian (fluent), Spanish (native).

Profiles

- LinkedIn: http://it.linkedin.com/pub/roberto-vera/30/a46/357/
- ResearchGate: http://www.researchgate.net/profile/Roberto_Vera_Alvarez
- Google Citations: http://scholar.google.com/citations?user=X9OKjuEAAAAJ&hl=en
- Skype user: r78v10a07
- Twitter: https://twitter.com/r78v10a07

Scientific Interests

My scientific interests include the development of complex computational systems for Bioinformatics. One of my aims is to integrate cloud computing technologies with bioinformatics tools and pipelines. I believe that sophisticated computational environments can help us to perform more reliable and accurate bioinformatics experiments. Due to my skills in both Bioinformatics and Computer Science, I can play a good role as a communication bridge between biologist and computer scientists.

Experience

1. C, C++, Java, PHP, HTML5, JavaScript, Bash scripting and SQL advancing programming skills applied to:
    - Development of web sites and servers for Bioinformatics.
    - Design and implementation of relational database systems to integrate biological data sources.
    - Next Generation sequencing problems.
    - Algorithm for the estimation of the small molecular solvation energy.
    - QSAR functions for the estimation of aqueous solubility.
    - Superposition of 3D patterns on chemical molecules.
    - Force Field programming (MMFF).
    - Scoring functions for receptor-ligand docking.
2. Advancing Linux Operating Systems Administration and Cloud computing:
    - Design, implementation and administration of a Beowulf computer cluster.
    - Computer network design and configuration.
    - Linux OS advancing programming and configuration.
    - Google Cloud Platform specialist.
3. Author of a fast NGS read aligner program designed for identifying pathogenic microbes in blood samples, that reached 19th position out of 2789 applicants in the Innocentive programming contest "Identify Organisms from a Stream of DNA Sequences" sponsored by the Defense Threat Reduction Agency under Air Force Contract #FA8721-05-C-0002.
4. Drug Design using Virtual Screening strategies. Design and validation of methodologies applied to several research projects.

Current projects

1. **JBioWH,** a Java based open-source framework for integration of biological data. https://code.google.com/p/jbiowh/

    Duties:

    - Design, implementation and maintenance of the whole project and web site.

    Languages and technologies:

    - Java
        i. Java Persistence API, EclipseLink Library.
        ii. JDBC, MySQL Connector.
        iii. Java GUI widget toolkit, Swing.
        iv. XML processing, SAX and DOM.
        v. RESTFul Web services, Jersey framework.
        vi. Context and dependency injection, Jboss Weld.
        vii. Java Server Faces, Glassfish and PrimeFaces.
        viii. Apache Commons for FTP Client.
        ix. Logging framework, SLF4J.

- Project management with Maven.
- Version Control system used: Subversion.

2. **Taxoner**, a simple, parallelizable workflow for taxonomic identification in NGS data. https://code.google.com/p/taxoner/

   Duties:

   - Design and implementation of the Gene assignment module. https://code.google.com/p/taxoner/wiki/05_GeneAssigment
   - Design and implementation of the graphical interfaces. https://code.google.com/p/taxoner/wiki/06_Graphical
   - Design and implementation of a Taxoner server. http://pongor.itk.ppke.hu/taxoner/
   - Design and implementation and maintenance of the project web site at Google Code.
   - Design and implementation of a Taxoner java based server on the Google Cloud Platform using the App Engine and the Compute Engine.

   Languages and technologies:

   - Java
     i. Context and dependency injection, Jboss Weld.
     ii. Java Server Faces, Glassfish and PrimeFaces.
     iii. Logging framework, SLF4J.
     iv. Google App Engine Java API.
   - C
     i. File and data structure indexing using B+Tree algorithm.
   - JavaScript, HTML5 and CSS3.
     i. NodeJS and AngularJS based web site.
   - Project management with Maven and Makefile.
   - Version Control system used: Subversion.

3. **BioC**, a C library for Bioinformatics. https://code.google.com/p/bioc/

   Duties:

   - Design, implementation and maintenance of the whole project and web site.

   Languages and technologies:

   - C
     i. File and data structure indexing using B+Tree algorithm.
   - Project management with Makefile.
   - Version Control system used: Subversion.

4. **nCService**, a simple NodeJS based application that allows external applications execute programs in the server via URL. https://code.google.com/p/ncservice/

   Duties:

   - Design, implementation and maintenance of the whole project and web site.

   Languages and technologies:

   - JavaScript, HTML5 and CSS3.
     i. NodeJS and AngularJS based web site.
   - Version Control system used: Subversion.

5. A java based web site running on the Google Cloud Platform using the App Engine. http://1.bioinformatics-platform.appspot.com/

   Duties:

   - Design, implementation and maintenance of the whole source code and web site.

   Languages and technologies:

   - Java
     i. Context and dependency injection, Jboss Weld.
     ii. Java Server Faces, Glassfish and PrimeFaces.
     iii. Logging framework, SLF4J.
     iv. Google App Engine Java API.
   - Project management with Maven.
   - Version Control system used: Subversion.

6. A virtual server for running bioinformatics software on the Google Cloud Platform. http://23.251.130.64/

   Duties:

   - Installation, configuration and maintenance of a Drupal based web site.
   - Installation, configuration and maintenance of the virtual machines.
   - Installation, configuration and maintenance of bioinformatics software.

   Languages and technologies:

   - Linux advanced administration.
     i. Configuration and installation of virtual machines.
     ii. Bioinformatics software compilation, configuration and installation.
     iii. Apache HTTP server configuration.
     iv. Apache Tomcat server configuration.
     v. NIS, NFS, AutoFS server configuration.
   - Drupal installation and configuration.
   - Google Cloud Platform administration.

List of Publications

1. Hudaiberdiev S, Choudhary S, **Vera Alvarez R**, Gelencsér Z, Ligeti B, Lamba D, Pongor S. Census of solo LuxR genes in prokaryotic genomes. *Frontiers in Cellular and Infection Microbiology*. 2015; **5**(4). DOI: 10.3389/fcimb.2015.00020.

2. **Vera R,** Pongor LS, Ligeti B. Fast and Sensitive Alignment of Microbial Whole Genome Sequencing Reads to Large Sequence Datasets on a Desktop PC: Application to Metagenomic Datasets and Pathogen Identification. *PLoS ONE*. 2014; **9**(7): e103441.

3. Dogsa I, Choudhary KS, Marsetic Z, Hudaiberdiev S, **Vera R**, Pongor S., Mandic-Mulec I. ComQXPA Quorum Sensing Systems May Not Be Unique to *Bacillus subtilis*: A Census in Prokaryotic Genomes. *PLoS ONE*. 2014; **9**(5): e96122.

4. Ligeti B., **Vera R.**, Gergely L., Balazs G., Pongor S. Predicting effective drug combinations via network propagation. *Biomedical Circuits and Systems Conference (BioCAS)*, 2013 IEEE. DOI: 10.1109/BioCAS.2013.6679718.

5. Kertész-Farkas A., Reiz B., **Vera R.**, Myers M. P. and Pongor S. PTMTreeSearch: a Novel Two-Stage Tree Search Algorithm with Pruning Rules for the Identification of Post-Translational

Modification of Proteins in MS/MS Spectra. *Bioinformatics*. 2014; **30**(2): 234-41. DOI: 10.1093/bioinformatics/btt642.

6. **Vera R.**, Perez-Riverol Y., Perez S., Ligeti B., Kertész-Farkas A., Pongor S. JBioWH: an open-source Java framework for bioinformatics data integration. *Database The Journal of Biological Databases and Curation.* 2013. DOI:10.1093/database/bat051.
7. **Vera R**, Perez-Riverol Y, Mazola Y, Mussachio A. A Parallel Systematic-Monte Carlo Algorithm for exploring Conformational Space. *Current topics in medicinal chemistry*. 2012; **12**(16): 1790-1796.
8. Perez-Riverol Y, Sánchez A, Ramos Y, Schmidt A, Müller M, Betancourt L, González LJ, **Vera R**, Padron G, Besada V. In silico analysis of accurate proteomics, complemented by selective isolation of peptides. *Journal of proteomics*. 2011; **74**(10):2071-82.
9. Sanchez A, Perez-Riverol Y, Gonzalez LJ, Noda J, Betancourt L, Ramos Y, Gil J, **Vera R**, Padron G, Besada V. Evaluation of Phenylthiocarbamoyl-Derivatized Peptides by Electrospray Ionization Mass Spectrometry: Selective Isolation and Analysis of Modified Multiply Charged Peptides for Liquid Chromatography-Tandem Mass Spectrometry Experiments. *Anal. Chem.* 2010; **82**: 8492–8501.

List of Patents

1. Chemical compounds having antiviral activity against dengue virus and other flaviviruses, WO/2009/106019
2. Antineoplastic compounds and pharmaceutical compositions thereof, WO/2006/119713.

Conferences

1. Biotechnology Havana'09. Havana 2009.
   - Mazola Y., Chinea G., Guirola O., **Vera R.**, Huerta V., Fleitas N., Mussachio A. Application of computational docking approach targeting the enveloped glycoprotein of Dengue virus.
   - Guirola O., Mazola Y., **Vera R.**, Perez-Riverol Y., Chinea G. Evaluation of docking programs for computer-assisted drug design.
   - Pérez-Riverol Y., **Vera R.**, Mazola Y., Millán A. SintCompound: An Smal database for Virtual Screening.
2. XXXII Congress of Theoretical Chemists of Latin Expression, Havana 2007.
   - **Vera R.** A new Systematic-Monte Carlo Algorithm to explore the molecule conformational space. (poster)
3. XII International Convention and Informatics Fair, Havana 2007.
   - **Vera R.** Estimating the aqueous solubility of small compounds. (Oral presentation)
4. I International Bioinformatics Congress, Havana, 2004.
   - **Vera R**., Batlle R., Lopez N., Rodríguez R. Estimation of the ligand binding energy for virtual screening. BIO046 pp. 386. (poster)
   - Batlle R., Lopez N., **Vera R.**, Rodriguez R. Distributing massive docking experiments for virtual screening. BIO050 pp. 388. (poster)
   - Barberena K., Lopez N., **Vera R.**, Batlle R., Rodríguez R. A relational database of chemical ligands for QSAR and virtual screening. BIO047 PP. 386. (poster)
5. II International Open Source Symposium, Havana 2004.

- **Vera R.**, Rodríguez R., Batlle R., Lopez N. Complex System simulation in Linux Clusters. SWL005 pp. 406. (poster)
- Rodriguez R., Lopez N., Batlle R., Guirola O., **Vera R.**, Borroto C., Alvarez F. Data mining and biomolecular simulations on Linux Clusters. SWL005 pp. 403. (poster)
- Borroto C., **Vera R.**, Rodríguez R., López N., Batlle R., Guirola O., Castro E. Open Source tools for multidisciplinary teems work. SWL005 pp. 404.