

Postdoc fellowship in Data Science / Bioinformatics (with application to immunology), UIC, Barcelona, Spain

Objective

Applications are invited for a Postdoctoral fellowship working under the supervision of Bernhard Knapp in the Department of Basic Sciences, Faculty of Medicine and Health Sciences, UIC University, Barcelona, Spain.

The work group of Bernhard Knapp (<https://www.uic.es/bioinformatics>) focuses on computational simulations of the (human) immune system. We provide novel insight in fundamental processes of the immune system by means of quantitative analysis and predictive models. More specifically the work focuses on how T-cell receptors (TCRs) recognise Major Histocompatibility Complex (MHC) bound peptides in different health and disease conditions as for example allergies, cancer, autoimmune diseases, or infections. Apart from TCR/peptide/MHC interactions also antibody/antigen interactions are a key research interest of the group.

We apply a wide range of computational data science methods for (biomedical) challenges. Examples of work include: by far the largest dataset of T-cell receptor recognition processes (Knapp et al., 2014), comparison of the molecular design (Dunbar et al., 2014) and instability (Knapp et al., 2017) of antibodies with T-cell receptors, obtaining novel insights in peptide/MHC detachment processes (Knapp et al., 2016), giving accurate and reproducible predictions of peptide/MHC binding affinities (Wan et al., 2015), optimising binding affinities using Genetic Algorithms (Knapp et al., 2011), and investigating protein/protein interactions (Esmailbeiki et al., 2016), and development of tools (Knapp et al., 2018). Highly parallel calculations are carried out on different supercomputers.

Requirements

Must criteria:

- **PhD** in a quantitative field (computer science, statistics, math, data science, (bio)physics, bioinformatics) earned within the **last 10 years**
- **BSc/MSc** in a quantitative field (e.g. computer science, statistics, math, data science, (bio)physics, bioinformatics).
- Research experience of at least **1 year outside of Spain** (postdoc experience abroad preferred)
- One or more **publications** in an international peer reviewed journal as **first-author** (for computer science candidates also peer reviewed conference papers are accepted)
- **Broad experience** and fluency in several **programming** languages such as Python, Matlab, C++, Java, Perl, and/or R.
- Strong **statistics** knowledge (univariate, multivariate)
- Strong **analytical skills, creativity, curiosity, enthusiasm**, and ability to work in a **team**
- Excellent command of the **English** language

Ideal criteria:

- Basic knowledge of molecular biology, genetics, immunology is an advantage but not required
- Previous experience with Monte Carlo simulations is an advantage

- Knowledge of data science techniques: Regression (e.g. ANN, Lasso, Ridge) , classification (e.g. SVM, KNN), clustering (e.g. k-means, hierarchical), heuristic optimisation (e.g. genetic algorithms), parallel computing (e.g. MPI, pySpark), SQL, noSQL is an advantage
- Knowledge of common bioinformatics techniques as Immunoinformatics, B- and T-cell epitope predictions, molecular dynamics simulations, free energy predictions, protein/ligand and protein/protein docking, virtual screening, protein structure prediction, sequence alignments, tree building is an advantage
- Knowledge of LINUX operating systems and command line operations is an advantage
- Experience with the use of high performance computing clusters is an advantage
- Previous teaching experience in English and/or Spanish is an advantage

Characteristics of the fellowship

- The fellowship is for 2 years. The expected start date is 1 October 2018.
- The researcher will have 7.000€ per year for research expenses. These expenses may include: consumables, travels, publication costs, congresses, etc.
- Access to a high performance computational cluster is available.
- Under no circumstances does this postdoctoral research contract imply a commitment by the University to subsequently hire the beneficiary as a permanent member of the University's staff.

Please read the application guidelines carefully and follow them as applications ignoring the guidelines cannot be considered:

Application documents

- 1) CV (including date of PhD degree and dates of research experience outside of Spain)
- 2) Publication list (including journal impact factors)
- 3) Letter of motivation (short, max 1 page) answering to the “must/ideal criteria” as stated above and please also state how you have heard about this position.
- 4) Contact details of 2-3 references

In a first step the applicants are encouraged to pass a quick test in order to decode the application email. Please apply to:

[1][2][3][4][5][6]@uic.es

where answering the questions underneath decodes the characters 1 to 6 above (without the square brackets and lower case). (Hint: You know that you have decoded correctly if the email address has to do with the name of the head of the work group)

[1] A variable that can only be true or false (only the first character)

- [2] Simplest and most common clustering algorithm (only the first character before the “-“)
- [3] A pointer to nothing in C/C++ (only the first character; not “void”)
- [4] Building blocks of proteins (only the first character)
- [5] Probability for a given statistical model that (if the null hypothesis is true) that the statistical summary (e.g. mean difference between two compared groups) would be the same as or more extreme than the actual observed results (only the first character before the “-“)

[6] import numpy as np

```
idx = np.argmax([99, 9, 1, 98, 7])
```

```
print("python!"[idx])
```

Please apply latest by the 15th of June 2018.

Reference List

- Dunbar,J., Knapp,B., Fuchs,A., Shi,J., and Deane,C.M. (2014). Examining Variable Domain Orientations in Antigen Receptors Gives Insight into TCR-Like Antibody Design. *PLoS Comput Biol* *10*, e1003852.
- Esmailbeiki,R., Krawczyk,K., Knapp,B., Nebel,J., and Deane,C.M. (2016). Progress and Challenges in Predicting Protein-Protein Interfaces. *Brief Bioinform* *17*, 117-31.
- Knapp,B., Alcala,M., Zhang,H., West,C., van der Merwe,P.A., and Deane,C.M. (2018). pyHVis3D: Visualising Molecular Simulation deduced H-bond networks in 3D: Application to T-cell receptor interactions. *Bioinformatics*.
- Knapp,B., Demharter,S., Deane,C.M., and Minary,P. (2016). Exploring peptide/MHC detachment processes using Hierarchical Natural Move Monte Carlo. *Bioinformatics* *32*, 181-186.
- Knapp,B., Dunbar,J., Alcala,M., and Deane,C.M. (2017). Variable Regions of Antibodies and T-cell Receptors may not be Sufficient in Molecular Simulations Investigating Binding. *J. Chem. Theory Comput.* *accepted*.
- Knapp,B., Dunbar,J., and Deane,C.M. (2014). Large Scale Characterization of the LC13 TCR and HLA-B8 Structural Landscape in Reaction to 172 Altered Peptide Ligands: A Molecular Dynamics Simulation Study. *PLoS Comput Biol* *10*, e1003748.
- Knapp,B., Giczi,V., Ribarics,R., and Schreiner,W. (2011). PeptX: Using Genetic Algorithms to optimize peptides for MHC binding. *BMC. Bioinformatics.* *12*, 241.
- Wan,S., Knapp,B., Wright,D., Deane,C., and Coveney,P.V. (2015). Rapid, Precise and Reproducible Prediction of Peptide-MHC Binding Affinities from Molecular Dynamics that Correlate Well with Experiment. *J. Chem. Theory Comput.* *11*, 3346-3356.