Leids Universitair Medisch Centrum

Bioinformatics Internship master student

Description

For our research group 'Immune Monitoring' at the department Immunohematology and Blood Transfusion we are looking for a Master student in medical science or Bioinformatics for an internship project (6-8 months) with basic programming skills in R or Python.

This concerns an internship in the English language.

Information

Evolutionary insights into T cell receptor germline alleles.

Indu Khatri12, Magda Berkowska[1], Erik van den Akker[2], Marcel Reinders[2] and Jacques van Dongen[1]

Affiliations:

[1]Department of immunohematology and Blood Transfusion
[2]Leiden Computational Biology Center

T cell receptors are composed of an alpha chain and a beta chain. The T cell receptor genes contain multiple V, D and J gene segments in their beta chains (and V and J gene segments in their alpha chains) that are rearranged during the development of the lymphocyte to provide that cell with a unique antigen receptor. The T cell receptor in this sense is the topological equivalent to an antigen-binding fragment of the antibody, both being part of the immunoglobulin superfamily.

Several databases viz. IMGT, IgPdb, VBASE2 report the germline variations in different individuals but IMGT is the most widely used database. The current database is more sampled from European populations and completely ignore the occurrence of alleles in other populations. This brings in an opportunity to look for the variations in all the populations and to understand what every population comprises in terms of immunoglobulin germline variation. Some of the estimates for identification of population specific variations can already be foreseen with the big genome projects. The completeness and accuracy of germline sequences may greatly influence downstream analysis in repertoire sequencing (Rep-Seq) studies where unreported alleles can appear as recurrent somatic mutation and skew estimated segment distributions in clinically relevant decision processes e.g. infection response and vaccination studies.

To address the questions from the perspective of existing germline variations in different populations we will use 1000 Genomes dataset. We will frame a set of rules to obtain the alleles with none sequencing error or SHM identified as germline allele from this data resource. Further, the complete set of alleles from different populations would let us understand if the differences in alleles are shaped by environmental pressure (i.e. random) or an effect of human migrations.

Contact

If you want to apply, please send your English CV to dr.Indu Khatri (i.khatri@lumc.nl).
You can also contact her for more information.