Computer aided data analysis at the Institute of Immunology

Virology

The Institute of Immunology has a long lasting experience in the analysis of human and veterinary virus genomes. In general complete or partial virus genomes are amplified by (RT-) PCR and nucleotide sequences are determined by cycle sequencing. The corresponding sequences are analysed using bioinformatics in order to investigate the genetic diversity of viruses. Usually new sequences are first submitted to a BLAST search in order to retrieve closely related virus sequences in public databases. Sequence alignments, including previously published (reference) sequences as well as the new sequences, are computed to determine the genetic distances between the latter. On the basis of these alignments phylogenetic trees are calculated using appropriate algorithms, to illustrate the evolutionary relationship between viruses. The bioinformatics tools which are most commonly used for such applications include ClustalW, MEGA3, Seqscape, Phylip, and BioEdit.

Such studies were performed to support epidemiological investigations on virus transmission and helped to demonstrate local interruption of virus circulation, confirm epidemiological links between different outbreaks, distinguish between endemic and epidemic settings and to discriminate between vaccine and wild-type viruses. The genetic characterisation of viruses allowed to determine virus specific mutation rates, to describe major evolutionary events like recombinations or reassortments between different virus strains and to identify specific mutations which are related to the pathogenicity or drug resistance of viruses.

Psychoimmunobiology

Analysis of human genomic sequences is performed using GenomBench (Vector NTi), subsequently, *in silico* investigation of human promoters is performed either using our in-house Transfac server, or *In silico Phylogenetic Footprinting (ISPF)* software developed in house combining multi species Transfac predictions with genomic sequence alignments to provide an accurate refinement of the Transfac data. Routine molecular biology (sequence analysis, PCR primer design, restriction enzyme analysis) is performed using the vector NTi software package. Post sequencing variance is analysed for disturbance of the Hardy-Weinberg equilibrium and SNP linkage using the HaploView software.

CENTRE DE RECHERCHE PUBLIC

crp

SANTÉ

Molecular docking is performed with Autodock, and molecules visualised with DeepView, or PyMol. Advanced data analysis is performed using either the basic "R" environment, or the biology specific "BioConductor" developed from R.

Imaging

Investigating new technological applications that promote knowledge in Immunology has always been the Institute's top priorities. With the creation of the first Luxemburgish proteomics platform, 2003 marked the beginning of mass spectrometry in our laboratory. The FNR Nano project focused on imaging mass spectrometry using a CAMECA NanoSIMS instrument. It required the development of new methodologies, and introduced imaging at the Institute. Since then, several projects in psychobiology and virology rely on conventional fluorescence microscopy (confocal, apotome) and NanoSIMS microscopy to study protein expression in-vitro. The analysis of the resulting datasets is performed using the ImageJ imaging framework. The need for additional processing functions for image metadata extraction, ratio calculation and particle analysis motivated the development of novel tools for image data analysis, most of which are publicly available under various open-source licenses. Some of these utilities are considered a standard by the imaging community.

Proteomics and Mass spectrometry

The mass spectrometry core facility (MALDI-TOF TOF Ultraflex I from Bruker and Nano ESI-Ion Trap XCTplus from Agilent) at the Institute of Immunology can be used routinely for precise determination of peptide/protein molecular mass, sequence, as well as to identify contaminating impurities. Methods can be employed for identification of proteins which are separated on two dimensional gel electrophoresis (2D-DIGE) and processed with Decyder, a software dedicated to 2D-DIGE analysis. Using the MALDI, the peptide mass lists are generated after enzymatic digestion by Flexanalysis 2.4 and the requests on NCBI nr or SWISSProt databases are processed through our in-house MASCOT server. The resulting identifications are then imported into Biotools 3.0 for further spectra interpretations. MSMS data can be produced for isoforms characterisation, PTMs or de novo sequencing. For the Ion trap Mass spectrometer, the mass lists integrate data from MS to MSn spectra (DataAnalysis software), which are queried via our MASCOT server. Further queries are performed using Spectrum Mill, as reports generated by this software are essential to deeper characterisation of proteins.