**UNIVERSIDADE ESTADUAL DE SANTA CRUZ - UESC**

**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO - PROPP**

**DEPARTAMENTO DE CIÊNCIAS BIOLÓGICAS - DCB**

**PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR - PPGGBM**

### PROGRAMA DE DISCIPLINA

|  |  |
| --- | --- |
| **Código:** | CIB 080 |
| **Nome da disciplina:** | T.E. em GBM I: Statistical analyses in genetics and genomics |
| **Pré-requisitos:** |  |
| **Carga horária** | **Teórica:** 30 **Prática:** 0 **Total:** 30 |
| **Créditos:** | **Teórica:** 2 **Prática:** 0 **Total:** 2 |
| **Professor:** |  |
| **Assinatura:** |  |
| **Ementa:** | Nowadays, and whatever the domain (medicine, ecology, genetics,...), data analysis is primordial to improve our understanding of biological processes.. According to the type of data, different statistical analyses may be used, and selecting appropriate ones could be a difficult task Therefore, we propose in this module to study a set of statistical methods adapted to different types of data. Moreover an introduction will be done on the analyse of “big” data, as data generated by “omics” technologies. Relationship between datasets will be explored as well asvisualization of the correlation structure.  |
| **Objetivos:** | Students will be able to select and to perform adequate statistical analyses for analyzing data of different types by using the free software R. Another objective is to introduce multivariate approaches for analyzing one or more datasets. |
| **Metodologia:** | After a short tutorial on the freeware R, theoritical aspects of each statistical analysis will be studied. Then applications on real or simulated datasets will be done to implement these approaches.  |
| **Avaliação:** | Para a avaliação serão mensurados: participação nas aulas e atividades propostas, além da presença mínima de 75%. Serão aprovados aqueles que no conjunto dos critérios estabelecidos obtiverem um mínimo de 70% de rendimento. |
| **Conteúdo Programático:** | -Short tutorial on R-Descriptive statistics-Statistical tests-Linear regression-Analysis of variance-Multivariate analyses :  - Principal Component Analysis (PCA) - Partial Least Squares method (PLS) - Sparse versions of PCA and PLS - Partial Least Squares Discriminant Analysis (PLS-DA) |
| **Referências Bibliográficas:** | 1- Lê Cao K-A., Rossouw D., Robert-Granié C. and Besse P. 2008. A Sparse PLS for Variable Selection when Integrating Omics data. SAGMB 7(1)2- Lê Cao K.-A., Martin P.G.P, Robert-Granié C. and Besse, P. 2009. Sparse Canonical Methods for Biological Data3- Gonzalez, I., Lê Cao, K.-A., Davis, M. J., Dejean, S., et al. 2012. Visualising associations between paired omics data sets. BioData mining, 5(1), 19.4- Team, R. D. C. 2011. R 2.14. 1. |
|  |