

Disciplinas do 2º semestre/2023 e da 1ª e 2ª metade do 2º semestre/2023

NG245 - SEMINÁRIOS DO CURSO PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR - TURMA PMV

Tema: Ciclo de Seminários da Pós-Graduação do Instituto de Biologia

Créditos: 2

Horário: Quartas-feiras, das 11:00 às 13:00

Local/Sala: **a definir**

Período de oferecimento: Todo o 1º semestre (de 02/08/2023 a 06/12/2023)

Vagas: 100

Mínimo de alunos: 1

Responsável: **Pedro Manoel Mendes de Moraes Vieira** - pmvieira@unicamp.br

Estudantes especiais: aceita - solicitar autorização do professor responsável e seguir [instruções](#)

PROGRAMA: A ser definido.

CRONOGRAMA: Seminários semanais.

BIBLIOGRAFIA: A ser disponibilizada no período do oferecimento da disciplina.

NG262 - GENÉTICA DE POPULAÇÕES - TURMA MIZ

Créditos: 4

Horário: Segundas-feiras e Quartas-feiras, das 09:00 às 12:00

Local/Sala: **a definir**

Período de oferecimento: 2ª metade do 2º semestre (de acordo com o cronograma)

Vagas: 15

Mínimo de alunos: 5

Responsável: **Maria Imaculada Zucchi**

Estudantes especiais: Não aceita

PROGRAMA:

PARTE TEÓRICA

A disciplina tem como objetivo principal apresentar e discutir os fundamentos teóricos da Genética de Populações, suas aplicações para a Biologia da Conservação e para o Melhoramento de Plantas, e atual transição para Genômica Populacional e técnicas modernas de análise. Espera-se proporcionar aos alunos uma visão moderna dos mecanismos envolvidos na dinâmica dos genes nas populações naturais e artificiais, bem como das ferramentas atuais de análise destes processos. Do ponto de vista teórico, o conteúdo da disciplina abrange os seguintes tópicos: 1) Introdução à Genética de Populações; 2) Constituição genética de populações (frequências genotípicas e gênicas); 3) Marcadores moleculares e seu uso na análise genética de populações; 4) Equilíbrios genéticos, Hardy-Weinberg; 5) Equilíbrio de ligação; 6) Equilíbrio de Wright; 7) Sistemas reprodutivos e de cruzamento; 8) Deriva genética; 9) Estrutura genética de populações naturais e artificiais (estatísticas-F de Wright e correlatas); 10) Equilíbrio mutação-deriva; 11) Tamanho efetivo populacional; 12) Metodologias de sequenciamento de nova geração; 13) Genômica de populações.

PARTE PRÁTICA

Utilização de programas computacionais em aulas práticas para estimação de parâmetros populacionais e análises especializadas com base em dados de marcadores moleculares de segunda e nova geração. Exemplos de programas computacionais a serem utilizados: GenAlex, MLTRwin, Structure, DARwin, pacotes populacionais do Rstudio (R) – adegenet, hierfstat, poppr, entre outros.

CRONOGRAMA:

14/11/2023 A 30/11/2023

BIBLIOGRAFIA:

Alendorf F., Luikart G., Aitken N. S. Conservation and the Genetics of Populations, Sec. Ed. Blackwell Publishing, 2013.
Falconer, D.S. & Macky, T. Introduction to Quantitative Genetics. 4 Ed. Longman Scientific & Technical, London, 1996.
Falconer, D.S. Quantitative genetics, 2nd Ed. Logman, News York.1981
Gondro, C. 2015. Primer to analysis of genomic data using R, Springer, 1 Ed. 2015.
Hartl, D.L. & Clark A.G. Principles of Populations Genetics. Second edition. Sinauer Associates, Inc. 1997.
Nei, M. Molecular evolutionary genetics. Columbia University Press. 1987.
Weir, S. B. Genetics Data Analysis II: Methods for Discrete Population Genetic Data, 1996.

NG281 - TÓPICOS AVANÇADOS DO PPG-GBM II - TURMA RFL

Tema: Recentes avanços em Bacteriologia Molecular

Créditos: 2

Horário: Segundas-feiras, das 10:00 às 12:00

Local/Sala: **a definir**

Período de oferecimento: Todo o 1º semestre (de 31/07/2023 a 04/12/2023)

Vagas: 15

Mínimo de alunos: 5

Responsável: **Rogério Ferreira Lourenço**

Colaboradora: **Cristina Elisa Alvarez Martinez**

Estudantes especiais: Não aceita

PROGRAMA:

Temas envolvendo genética e biologia celular de sistemas bacterianos recém-publicados em revistas de Microbiologia. Serão usados artigos que tenham sido depositados no BioRxiv como *preprint*.

CRONOGRAMA:

Aula 1: Introdução à disciplina; Aulas 2, 4, 6, 8, 10, 12 e 14: Apresentação e discussão aberta de artigo; Aulas 3, 5, 7, 9, 11, 13 e 15: Discussão dirigida e comparação do conteúdo do *preprint* com o do artigo aceito.

BIBLIOGRAFIA:

A ser disponibilizada no período do oferecimento da disciplina.

NG282 - TÓPICOS AVANÇADOS DO PPG-GBM III - TURMA RVS

Tema: Análise de Dados Biológicos em R

Créditos: 3

Horário: Quartas-feiras, das 15:00 às 18:00

Local/Sala: **Prédio do LGE.**

Período de oferecimento: Todo o 1º semestre (de 02/08/2023 a 06/12/2023)

Vagas: 30

Mínimo de alunos: 10

Responsável: **Renato Vicentini dos Santos**

Colaboradores: **Luana Walravens Bergamo e Lucas Miguel de Carvalho**

Estudantes especiais: Não aceita

PROGRAMA:

A disciplina abordará conceitos básicos e avançados em análise de dados utilizando a Linguagem R. Ela é destinada principalmente a alunos que tenham a intenção de analisar dados qualitativos e quantitativos como ferramenta de estudo em seus projetos de pesquisa, mesmo aqueles que não tenham conhecimento em programação. Ao final dessa disciplina, esperamos que o aluno seja capaz de analisar seus próprios dados, aplicando testes estatísticos, produzindo gráficos simples e complexos, aplicando análises de classificação e produzindo relatórios de dados. Os temas a serem abordados serão:

1. Lógica Computacional
2. R básico
3. Estatística básica
4. Dataframes e arquivos
5. Manipulações de datasets
6. Limpeza de dados
7. Funções
8. Dplyr e Tidyverse
9. Ggplot
10. Teste t
11. ANOVA
12. Testes não paramétricos
13. Teste qui-quadrado
14. Regressões lineares
15. Principal Component Analysis (PCA)
16. Heatmaps
17. Clustering – Parte 1
18. Clustering – Parte 2
19. Árvore de decisão
20. Curvas ROC
21. Árvore de regressão
22. Mapas
23. Distribuição hipergeométrica e ajuste de p-values
24. Relatório de dados
25. Análises de genética populacional

CRONOGRAMA: Será disponibilizado no início da disciplina.

BIBLIOGRAFIA:

Wickham H, Golemund G. R for Data Science. O'Reilly; 2017.

Kabacoff RI. R in action: Data analysis and graphics with R. Manning; 2015. Teetor P. R Cookbook. O'Reilly; 2011. Bivand RS, Pebesma E, Gómez-Rubio V. Applied spatial data analysis with R. Springer; 2013.

STHDA: Statistical tools for high-throughput data analysis. Disponível em <http://www.sthda.com/english/>

Tutorials point: simply easy learning. Disponível em: <https://www.tutorialspoint.com/r/index.htm>.

Lantz, Brett. Machine Learning with R - Second Edition; Packt Publishing; 2015

NG283 - TÓPICOS AVANÇADOS DO PPG-GBM IV - MIZ

Tema: Construção de Bibliotecas GBS (Genotyping By Sequencing) visando a descoberta de SNPs

Créditos: 4

Horário: Segundas-feiras, Terças-feiras e Quartas-feiras, das 09:00 às 17:00

Local/Sala: **Barracão da Genética**

Período de oferecimento: 1ª metade do 2º semestre (de acordo com o cronograma)

Vagas: 10

Mínimo de alunos: 5

Responsável: **Maria Imaculada Zucchi**

Colaboradora: **Anete Pereira de Souza**

Estudantes especiais: Não aceita

PROGRAMA:

O objetivo deste curso é aprender passo a passo como preparar bibliotecas para Genotyping By Sequencição (GBS). No final do curso, os alunos deverão ser capazes de:

- 1) Descrever os passos necessários para preparar bibliotecas para GBS;
- 2) Avaliar a qualidade de bibliotecas de DNA digeridas com diferentes enzimas de restrição;
- 3) Preparar uma biblioteca GBS para ser enviada para sequencição Illumina. Iremos discutir também algumas das possíveis aplicações dos dados gerados por GBS em estudos de genômica populacional;
- 4) Mostraremos abordagens para filtragem de dados e identificação de marcadores SNPs.

Este curso abordará aspectos teóricos e práticos do desenvolvimento de bibliotecas genômicas GBS visando a identificação de marcadores SNP, filtragem de dados usando o pipeline de Stacks e Ipyrad, e introdução à análise genômica de populações.

CRONOGRAMA:

28 de agosto a 01 de setembro

BIBLIOGRAFIA:

A ser disponibilizada no período do oferecimento da disciplina.

NG296 - ESCRITA ACADÊMICA - TURMA MFN

Obs.: A disciplina será ministrada em conjunto com a NF140/MBJ - Elaboração de Artigos Científicos (PPG-BMM)

Créditos: 4

Horário: Quintas-feiras, das 14:00 às 17:00

Local/Sala: **a definir**

Período de oferecimento: Todo o 1º semestre (de acordo com o cronograma)

Vagas: 15

Mínimo de alunos: 5

Responsável: **Mariana Freitas Nery**

Estudantes especiais: Não aceita

PROGRAMA:

The course is based on activities done in class and homework. Proofreading is not part of this discipline; instead, students will plan, write, and revise theirs and peers' manuscripts. We aim to improve students' writing skills by raising their awareness about text organization, flow, and coherence. Students will build up their toolkit to self-improve their writing skills after the course.

CRONOGRAMA:

Week 1: Need Analysis, Outline of the Course, Characteristics of Scientific Writing

Week 2: Sentence 03/08

Week 3: Paragraph 10/08

Week 4: Mind mapping 17/08

Week 5: Title & Abstract 24/08

Week 6: Introduction 31/08

Week 7: Methods 13/09

Week 8: Figures, Tables, Legends 20/09

Week 9: Results 27/09

Week 10: Discussion 05/10

Week 11: Conclusions & Acknowledgement 19/10

Week 12: Peer review of students writing 26/10

Week 13: Peer analysis (S-S) and improving your writing 02/11

BIBLIOGRAFIA:

Reading suggestions

Gastel B., Day R. A. (2016) "How to Write and Publish a Scientific Paper", 8th Edition. Greenwood.●

Schimel J. (2012) "Writing Science: How to Write Papers That Get Cited and Proposals That Get Funded". Oxford: Oxford University Press.

Schuster, Ethel, Levkowitz, Haim, Oliveira Jr., Osvaldo N. (Eds) (2014) "Writing Scientific Papers in English Successfully: Your complete roadmap", Gráfica Compacta.●

Other materials will be suggested or made available (via Moodle) throughout the course.

NG300 - DO GENE AO FENÓTIPO: UMA INTRODUÇÃO AO PIPELINE DE BIOTECNOLOGIA AVANÇADA - TURMA PA

OBS.: Interessados além de se inscreverem no sistema de matrícula de disciplinas, também deverão se inscrever no site: <https://www.gccrc.unicamp.br/course-from-gene-to-trait/>

Créditos: 3

Horário: (de acordo com o programa e cronograma)

Local/Sala: **Auditório Embrapa Agricultura Digital**

Período de oferecimento: 1ª metade do 2º semestre (de acordo com o programa e cronograma)

Vagas: 100

Mínimo de alunos: 10

Responsável: **Paulo Arruda** - parruda@unicamp.br

Estudantes especiais: aceita - solicitar autorização do professor responsável e seguir [instruções](#)

PROGRAMA:

September 18th- Monday

08:30 - 10:00 Welcome: The GCCRC “From Gene to Trait” platform - Paulo Arruda (UMIP/GCCRC Unicamp) and Stanley Oliveira (General Head of Embrapa Digital Agriculture)

10:30 - 12:00 Overview on gene discovery: an industry biotech program to enhance maize agronomic traits Carl Simmons (Corteva Agriscience) (TBC)

13:30 - 15:00 Palestra online Todd Michael (TBC)

15:30 - 17:00 Gene discovery in wild species for abiotic stress tolerance - Isabel Gerhardt – (UMIP/GCCRC Embrapa)

September 19th- Tuesday

08:30 - 10:00 Plant microbiome for genes and traits - Rafael Souza ou Jader Armanhi (Symbiomics)

10:30 - 12:00 Título a definir – Silvia Souza (TBC)

13:30 - 15:00 Título a definir – Adriana Hemmerly (TBC)

15:30 - 17:00 The Brazilian case of biological nitrogen fixation in sustainable crop production - Solon Araujo (SCA Consultoria) –

September 20th- Wednesday

08:30 - 10:00 CRISPR-Cas9 genome editing in maize for sustainable Agriculture - José H. Lopes Filho (UMIP/GCCRC, Embrapa)

10:30 - 12:00 Soja editada – Liliane Hennig (Embrapa SOja)

13:30 - 15:00 Protoplasma como modelo – Sofya Gerasimova (UMIP/GCCRC Embrapa)

15:30 - 17:00 Designing strategies for vectors and constructs – Almir Zanca (Stoller Brasil)

September 21th- Thursday

08:30 - 10:00 Multiplex gene editing in maize- Laurens Pauwels (VIB-UGent, Belgium) - online

10:30 - 12:00 Título a definir - Fátima Grossi (Cenargen Embrapa)

13:30 - 15:00 Trait introgression in commercial maize hybrids - Ivan Schuster (Long Ping Sementes) - TBC

15:30 - 17:00 Fenotipagem (Título a definir) - Juliana Yassitepe (UMIP/GCCRC Embrapa)

September 22th- Friday

08:30 - 10:00 Dereglulation of biotechnological traits – Adriana Chiavegato (CTC)

10:30 - 12:00 Experiências e dificuldades de produzir produtos transgênicos – Adriana Capela (Oxitec)

13:30 - 15:00 Intellectual property of biotechnological traits – Priscila Kashiwabara (TBC)

15:30 - 17:00 Final considerations and concluding remarks - Paulo Arruda (UMIP/GCCRC Unicamp)

CRONOGRAMA:

Disciplina CONDENSADA será ministrada do dia 18 a 22 de setembro de 2023 no horário de 8h-17h, no auditório da Embrapa Agricultura Digital

Av. Dr. André Tosello, 209 - Cidade Universitária, Campinas - SP

BIBLIOGRAFIA:

A disciplina será ministrada por pesquisadores nacionais e internacionais associados ao GCCRC (Genomics for Climate Change Research Center), vinculado a UMIPGenClima, Unidade Mista de Pesquisa Embrapa Unicamp.

NG302 – BIOLOGIA SISTÊMICA E INTEGRATIVA APLICADA NA ANÁLISE DE DADOS MOLECULARES - TURMA MMB

Créditos: 3

Horário: Segundas-feiras e das Quartas-feiras, das 14:00 às 17:00

Local/Sala: **Auditório do CBMEG**

Período de oferecimento: 2ª metade do 2º semestre (de 02/10/2023 a 06/12/2023)

Vagas: 30

Mínimo de alunos: 8

Responsável: **Marcelo Mendes Brandão** - brandaom@unicamp.br

Estudantes especiais: aceita - solicitar autorização do professor responsável e seguir [instruções](#)

PROGRAMA:

O advento das tecnologias de sequenciamento de alto desempenho, também conhecido como de nova geração, revolucionou a aquisição de dados biológicos nos últimos anos. Isso permitiu obter um grande volume de informações, o que teve um impacto significativo no estudo de questões relacionadas ao dogma central da Biologia Molecular.

Como resultado, os projetos de pesquisa atuais têm incorporado questões mais complexas, com abordagens mais preditivas e descritivas, combinando informações de diferentes fontes de dados moleculares para estudar sistemas e componentes biológicos envolvidos nos projetos propostos.

A Biologia Computacional e Bioinformática têm acompanhado de perto essa revolução na aquisição de dados, desenvolvendo novas ferramentas estatísticas, matemáticas e computacionais capazes de lidar com esses grandes volumes de informações e correlacioná-los com diferentes bancos de dados biológicos. Essa disciplina tem como objetivo apresentar aos estudantes abordagens práticas de Biologia Computacional e Bioinformática, capacitando-os a realizar análises qualitativas de arquivos de sequenciamentos, propor montagens de transcritos de novo ou baseadas em modelos genômicos disponíveis, avaliar a expressão diferencial de genes/contigs entre diferentes situações e anotar a função dos transcritos envolvidos na análise.

Além de capacitar os alunos nas ferramentas disponíveis no sistema operacional GNU/LINUX, é importante prepará-los para a era da ciência assistida por ferramentas de inteligência artificial (IA). A IA já está sendo amplamente utilizada em pesquisas científicas, desde a análise de dados até a criação de modelos preditivos complexos. Os cientistas modernos precisam estar preparados para lidar com o grande volume de dados gerados pelas tecnologias de sequenciamento de alto desempenho e extrair informações relevantes a partir deles. A IA pode ajudar nesse processo, automatizando tarefas repetitivas e permitindo que os cientistas se concentrem em análises mais complexas.

Ao incorporar ferramentas de IA em seu repertório de habilidades, os alunos estarão melhor preparados para enfrentar os desafios da ciência moderna e contribuir para o avanço do conhecimento em Biologia Molecular.

CRONOGRAMA:

1. **Nivelamento e atualização das habilidades usando GNU/Linux**
2. **Aplicações em Bioinformática de ferramentas para o gerenciamento de ambientes de computação: sistemas de gerenciamento de pacotes e virtualização de containers**
3. **Construção de Bancos de dados**
4. **Acesso e consulta a Banco de dados públicos**
5. **Lógica de programação**
6. **Reprodutibilidade das análises**
7. **Expressão diferencial /QC**
8. **Anotação funcional e estrutural de sequências primárias**
9. **Enriquecimento funcional**

BIBLIOGRAFIA:

Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* 12(1): 59-60. Calvo B, Larranaga P, Lozano JA. 2007. Learning Bayesian classifiers from positive and unlabeled examples. *Pattern Recognition Letters* 28(16): 2375-2384. Chevreux B, Pfisterer T, Drescher B, Driesel AJ, Muller WE, Wetter T, Suhai S. 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res* 14(6): 1147-1159. Clarke K, Yang Y, Marsh R, Xie L, Zhang KK. 2013. Comparative analysis of de novo transcriptome assembly. *Sci China Life Sci* 56(2): 156-162. Dean J, Ghemawat S. 2004. MapReduce: simplified data processing on large clusters. In: *Proceedings of the 6th conference on Symposium on Operating Systems Design & Implementation - Volume 6*. San Francisco, CA: USENIX Association. 10-10. Eddy SR. 2011. Accelerated Profile HMM Searches. *PLoS Comput Biol* 7(10): e1002195. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer EL, Tate J, Punta M. 2014. Pfam: the protein families database. *Nucleic acids research* 42(Database issue): D222-230. Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28(23): 3150-3152. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Fulllength transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29(7): 644-652. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, Macmanes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 8(8): 1494-1512. He JZ, Zhang Y, Li X, Shi P. 2012. Learning naive Bayes classifiers from positive and unlabelled examples with uncertainty. *International Journal of Systems Science* 43(10): 1805-1825. Kadota K, Nishiyama T, Shimizu K. 2012. A normalization strategy for comparing tag count data. *Algorithms Mol Biol* 7(1): 5. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y. 2008. KEGG for linking genomes to life and the environment. *Nucleic acids research* 36(Database issue): D480-484. Kersey PJ, Staines DM, Lawson D, Kulesha E, Derwent P, Humphrey JC, Hughes DST, Keenan S, Kerhornou A, Koscielny G, Langridge N, McDowall MD, Megy K, Maheswari U, Nuhn M, Paulini M, Pedro H, Toneva I, Wilson D, Yates A, Birney E. 2011. Ensembl Genomes: an integrative resource for genome-scale data from non-vertebrate species. *Nucleic acids research* 40(D1): D91-D97. Koski LB, Gray MW, Lang BF, Burger G. 2005. AutoFACT: an automatic functional annotation and classification tool. *Bmc Bioinformatics* 6: 151. Kumar S, Blaxter ML. 2010. Comparing de novo assemblers for 454 transcriptome data. *BMC Genomics* 11: 571. Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *Bmc Bioinformatics* 12: 323. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadukumar, Buchner A, Lai T, Steppi S, Jobb G, Forster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, König A, Liss T, Lussmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, Schleifer KH. 2004. ARB: a software environment for sequence data. *Nucleic acids research* 32(4): 1363-1371. Manning CD, Raghavan P, Schütze H. 2008. *Introduction to information retrieval*. New York: Cambridge University Press. Mlecnik B, Scheideler M, Hackl H, Hartler J, Sanchez-Cabo F, Trajanoski Z. 2005. PathwayExplorer: web service for visualizing high-throughput expression data on biological pathways. *Nucleic acids research* 33(Web Server issue): W633-637. Moreno-Hagelsieb G, Latimer K. 2008. Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* 24(3): 319-324. R-Core-Team. 2013. R: A language and environment for statistical computing. *Book R: A language and environment for statistical computing*. Supek F, Bosnjak M, Skunca N, Smuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6(7): e21800. Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva EV. 2013. OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic acids research* 41(Database issue): D358-365. Zhou S, Liao R, Guan J. 2013. When cloud computing meets bioinformatics: a review. *J Bioinform Comput Biol* 11(5): 1330002. Zou Q, Li XB, Jiang WR, Lin ZY, Li GL, Chen K. 2013. Survey of MapReduce frame operation in bioinformatics. *Brief Bioinform*.