Applied Population Genetic Analysis of Microbes

PLP6905C section 07EE

3 credit hours

Special Notes:

This course requires no previous experience in population genetics. Basic population genetics concepts will be introduced, but not derived in this course. Students are expected to have a basic understanding of genetics.

This course is designed to address specific problems faced when analyzing microbial populations. Plant pathogens and other microbes often do not conform to the assumptions underlying population genetic analysis, for example sexual populations or diploid genomes. The emphasis of this course is on practical aspects of data collection and analysis.

Course Description:

In this course students will learn to use DNA sequence or genetic marker data to describe population genetic variation and infer evolutionary processes in microbes. The emphasis will be on plant pathogens and examples from the plant pathology literature, but will also be applicable to other microbial populations. Topics to be covered include: sampling strategies, marker types and their evolution, genealogical inference, defining population and geographic structure, and coalescent-based methods for inferring demographic processes (e.g. divergence, migration, and recombination/sex). Methods will be applied in weekly labs. Students will have the opportunity to analyze their own population genetic data sets.

Instructor:

Erica Goss

Room 2415 Fifield Hall
emgoss@ufl.edu
352-273-4650

Course Time/ Location:

Lecture: Fifield Hall 2318, Tuesday and Thursday, period 7, 1:55-2:45 pm.

Lab: Microbiology and Cell Science Computer Lab, Friday, periods 5-7, 11:45-2:45 pm.

Polycom: Course may be taken via polycom. Students off-site must have a computer on which they will install freely available software for the computer labs. Please discuss logistics with instructor prior to the first week of class.

Office Hours:

Office hours are Tuesdays and Thursdays 2:45-4:00 pm in 2415 Fifield Hall, or by appointment.

Course Objectives:

Sophisticated computational methods are increasingly being coupled with population genetic data to infer the demographic history and evolution of populations. The use of these methods in inferring population processes is particularly critical for pathogens whose ecology or epidemiology may not easily be observed. In this course, you will learn how to use population genetic data to address questions such as: Is a new disease or unusual isolate a new species? Is a plant pathogen reproducing sexually? Is there migration among habitats or regions? How many times was an invasive pathogen introduced? The course will specifically address concerns specific to plant pathogens and microbes, which often clonally reproduce or have mixed reproduction systems. This course will provide you with the skills to decide on a sampling design, chose a marker system that will allow you to answer your questions of interest, analyze your data with state of the art methods, and interpret the results. Importantly, you will also learn the assumptions, limitations, and appropriate use of these methods.

Required and Recommended Textbooks:

There is no required textbook for this course. All required material will be provided, see reading list on pp. 7-9.

Assessment:

This is an upper level course and you are expected to be taking the course to expand your knowledge and improve your research. The assigned exercises and projects are meant to help you gain experience using these methods and synthesize what you have learned, but they will also be used to evaluate the level of careful thought and effort that you are putting into the course.

Laboratory Exercises

55% of your final grade (5% each) will come from laboratory exercises. These will be written up as mini-manuscripts. Details will be provided during the first lab.

Exam

15% of your grade will come from a mid-term exam.

Final Project

A final project will count for 30% of your grade. The project will be graded based on a research paper and presentation to the class. Specific guidelines for the paper and presentation will be provided. The project will involve analysis and interpretation of a population genetic data set. These data can come from the student's own research or the instructor can provide data. Students should discuss project ideas with the instructor early in the course. Results will be presented in a research style talk during the last three classes of the term. The paper is due the first day of the exam period, April 28.

Make-up and Attendance Policy:

Attending course lectures and labs, completing required assignments on time, and attending office hours for extra help as needed is expected. In this course lectures build on each other. Slides for the lectures will be made available, but notes from missed lectures must be obtained from other students in the course. Please contact the instructor directly regarding any serious illness or

prolonged absence. In exceptional circumstances, labs may be made up on your own time. However, the software used, while all freely available for download, is not necessarily simple to use. If you do not attend lab you may find yourself having trouble completing the exercise. A make-up exam will be provided in the case of a valid and documented excuse.

Course Schedule:

The following is an outline of what we will cover each class. This schedule is subject to change, and the instructor will inform you of any changes. Classes held in the Microbiology and Cell Science computer lab are indicated with the prefix "Lab". All others will be held in Fifield Hall.

Jan.	10	Introduction and evolutionary framework
	12	Population genetic theory, part I
	13	Lab: Command line introduction/population simulations
	17	Genetic markers – RFLP, AFLP
	19	Genetic markers – Microsatellites
	20	Lab: Working with data files
	24	Genetic markers – DNA sequences and SNPs
	26	Measures of population genetic variation
	27	Lab: Diversity statistics
	31	Measures of population genetic variation, continued
Feb.	2	Sampling strategies
	3	Lab: Linkage disequilibrium
	7	Inferring trees, methods & assumptions
	9	Inferring trees & networks
	10	Lab: Tree building
	14	Population structure – species, populations, "strains"
	16	Population structure – subdivison/geographic structure
	17	Lab: Analysis of molecular variance, F-stats
	21	Population structure – Bayesian inference
	23	Population structure – model-free inference
	24	Lab: Structure
	28	Phylogeography of plant pathogens
March	1	Exam
	2	Lab: DAPC
		Spring Break
	13	Population genetic theory, part II
	15	Coalescent theory
	16	Lab: Coalescent simulations
	20	Coalescent methods
	22	Demographic models – introduction, population size
	23	Lab: IM
	27	Demographic models – divergence
	29	Demographic models – migration
	30	Lab: Migrate
April	3	Demographic models – migration
	5	Multilocus inference
	6	Lab: Project work
	10	Time-calibrated genealogies
	12	Recombination
	13	Lab: BEAST

Academic Honesty, Software Use, UF Counseling Services, Services for Students with Disabilities

In 1995 the UF student body enacted a new honor code and voluntarily committed itself to the highest standards of honesty and integrity. When students enroll at the university, they commit themselves to the standard drafted and enacted by students.

In adopting this honor code, the students of the University of Florida recognize that academic honesty and integrity are fundamental values of the university community. Students who enroll at the university commit to holding themselves and their peers to the high standard of honor required by the honor code. Any individual who becomes aware of a violation of the honor code is bound by honor to take corrective action. The quality of a University of Florida education is dependent upon community acceptance and enforcement of the honor code.

The Honor Pledge: We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honesty and integrity.

On all work submitted for credit by students at the university, the following pledge is either required or implied: **"On my honor, I have neither given nor received unauthorized aid in doing this assignment."**

The university requires all members of its community to be honest in all endeavors. A fundamental principle is that the whole process of learning and pursuit of knowledge is diminished by cheating, plagiarism and other acts of academic dishonesty. In addition, every dishonest act in the academic environment affects other students adversely, from the skewing of the grading curve to giving unfair advantage for honors or for professional or graduate school admission. Therefore, the university will take severe action against dishonest students. Similarly, measures will be taken against faculty, staff and administrators who practice dishonest or demeaning behavior.

Students should report any condition that facilitates dishonesty to the instructor, department chair, college dean, Student Honor Council, or Student Conduct and Conflict Resolution in the Dean of Students Office.

(Source: 2011-2012 Undergraduate Catalog)

It is assumed all work will be completed independently unless the assignment is defined as a group project, in writing by the instructor.

This policy will be vigorously upheld at all times in this course.

Software Use:

All faculty, staff and students of the university are required and expected to obey the laws and legal agreements governing software use. Failure to do so can lead to monetary damages and/or criminal penalties for the individual violator. Because such violations are also against university policies and

rules, disciplinary action will be taken as appropriate.

Campus Helping Resources

Students experiencing crises or personal problems that interfere with their general well-being are encouraged to utilize the university's counseling resources. The Counseling & Wellness Center provides confidential counseling services at no cost for currently enrolled students. Resources are available on campus for students having personal problems or lacking clear career or academic goals, which interfere with their academic performance.

- University Counseling & Wellness Center, 3190 Radio Road, 352-392-1575, www.counseling.ufl.edu/cwc/
    - Counseling Services
    - Groups and Workshops
    - Outreach and Consultation
    - Self-Help Library
    - Training Programs
    - Community Provider Database

- Career Resource Center, First Floor JWRU, 352-392-1601, www.crc.ufl.edu/

Students with Disabilities

The Disability Resource Center coordinates the needed accommodations of students with disabilities. This includes registering disabilities, recommending academic accommodations within the classroom, accessing special adaptive computer equipment, providing interpretation services and mediating faculty-student disability related issues.

0001 Reid Hall, 352-392-8565, www.dso.ufl.edu/drc/

Grade Points

In accordance with current University of Florida policy, grade points will be assigned as follows.

Letter Grade	Grade Points	Percentage
A	4.0	90 or above
A-	3.67	87-89
B+	3.33	84-86
B	3.0	80-83
B-	2.67	77-79
C+	2.33	74-76
C	2.0	70-73
C-	1.67	67-69
D+	1.33	64-66
D	1.0	60-63
D-	0.67	57-59
E	0.0	56 or below
WF	0.0	NA
I	0.0	NA

NG	0.0	NA
S-U	0.0	NA

Detailed and up-to-date information on grades and grading policies can be found at the UF registrar web site, **www.registrar.ufl.edu/catalog/policies/regulationgrades.html**

Related Courses:

University of Florida offers a number of courses that touch on various aspects of population genetics. These include various courses in Evolutionary Biology and Evolutionary Genetics offered in the Biology Department. This course is specifically designed introduce students to the analysis of microbial population genetic data. Related courses offered through other programs tend to focus on plants and animals (e.g. WIS 6934: Molecular Ecology) and specific applications, e.g. plant breeding or wildlife conservation biology. This course is also distinctive in its emphasis on hands-on application of methods. Other courses to consider are:

BOT 6935: Molecular Systematics
Molecular Systematics is a course focusing on systematics, the evolutionary relationships among species and higher taxa, using molecular tools. It is an excellent complement to this course. More advanced systematics courses are also available through the Botany program.

ZOO6927: Statistical Principles for the Biological Sciences
Statistical Principles for the Biological Sciences is a new course on the statistics underlying most of the methods that will be applied in this course. For a deeper understanding of these methods, this course is highly recommended.

Applied Population Genetic Analysis of Microbes

Reading list (tentative, subject to change)

Jan.	10	Introduction and evolutionary framework	Carbone & Kohn 2004, McDonald & Linde 2002
	12	Population genetic theory, part I	Gillespie Ch 1-2, Wakeley 2005
	13	<u>Lab</u>: Command line introduction/population simulations	Wakeley Ch 1
	17	Genetic markers – RFLP, AFLP	Schlotterer 2004
	19	Genetic markers – Microsatellites	Ellegren 2004
	20	<u>Lab</u>: Working with data files	Excoffier & Heckel 2005
	24	Genetic markers – DNA sequences and SNPs	Li Ch 1 & 3
	26	Measures of population genetic variation	Grunwald et al. 2003
	27	<u>Lab</u>: Diversity statistics	TBD
	31	Measures of population genetic variation, continued	Halkett et al. 2005
Feb.	2	Sampling strategies	Felsenstein 2006
	3	<u>Lab</u>: Linkage disequilibrium	TBD
	7	Inferring trees, methods & assumptions	Avise 1989, Phylogenetic Handbook Ch 1
	9	Inferring trees & networks	Huson & Bryant 2006
	10	<u>Lab</u>: Tree building	Phylogenetic Handbook Ch 6
	14	Population structure – species, populations, "strains"	Feil & Spratt 2001
	16	Population structure – subdivison/geographic structure	Peever et al. 2004
	17	<u>Lab</u>: Analysis of molecular variance, F-stats	TBD
	21	Population structure – Bayesian inference	Pritchard et al. 2000
	23	Population structure – model-free inference	Jombart et al. 2010
	24	<u>Lab</u>: Structure	Structure manual
	28	Phylogeography of plant pathogens	Knowles 2009
March	2	<u>Lab</u>: DAPC	DAPC docs
	13	Population genetic theory, part II	Wakeley Ch 3
	15	Coalescent theory	Nordborg 2001
	16	<u>Lab</u>: Coalescent simulations	SNAP Workbench tutorial
	20	Coalescent methods	Kuhner 2009
	22	Demographic models – introduction, population size	Stukenbrock et al. 2007, Carbone et al. 2004
	23	<u>Lab</u>: IM	IM docs
	27	Demographic models – divergence	Hey and Nielsen 2004
	29	Demographic models – migration	Beerli 2009
	30	<u>Lab</u>: Migrate	Migrate docs
April	3	Demographic models – migration	Beerli & Palczewski 2010

5	Multilocus inference	Douhan et al. 2008
10	Time-calibrated genealogies	Drummond et al. 2006, Fargette et al. 2008
12	Recombination	Fearnhead & Donnelly 2001
13	Lab: BEAST	Drummond & Rambaut 2007
17	Selection	Phylogenetic Handbook Ch 14

Beerli, P. (2009). How to use migrate or why are Markov chain Monte Carlo programs difficult to use? Population Genetics for Animal Conservation. G. Bertorelle, M. W. Bruford, H. C. Hauffe, A. Rizzoli and C. Vernesi. Cambridge, Cambridge University Press. **17:** 42-79.

Beerli, P. and J. Felsenstein (1999). Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics **152**(2): 763-773.

Carbone, I. and L. M. Kohn (2004). Inferring process from pattern in fungal population genetics. Applied Mycology and Biotechnology Series. D. K. Arora. Amsterdam, The Netherlands, Elsevier Science. **4:** 29-58.

Carbone, I., Y. Liu, et al. (2004). Recombination and migration of *Cryphonectria hypovirus 1* as inferred from gene genealogies and the coalescent. Genetics **166**: 1611-1629.

Douhan, G. W., M. E. Smith, et al. (2008). Multigene analysis suggests ecological speciation in the fungal pathogen *Claviceps purpurea*. Molecular Ecology **17**(9): 2276-2286.

Drummond, A. J. and A. Rambaut (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology **7**.

Drummond, A. J., S. Y. W. Ho, et al. (2006). Relaxed phylogenetics and dating with confidence. PLoS Biology **4**(5): e88.

Ellegren, H. (2004). Microsatellites: Simple sequences with complex evolution. Nature Reviews: Genetics **5**: 435-445.

Excoffier, L. and G. Heckel (2006). Computer programs for population genetics data analysis: a survival guide. Nature Reviews: Genetics **7**(10): 745-758.

Fargette, D., A. Pinel, et al. (2008). Rice Yellow Mottle Virus, an RNA Plant Virus, Evolves as Rapidly as Most RNA Animal Viruses. Journal of Virology **82**(7): 3584-3589.

Fearnhead, P. and P. Donnelly (2001). Estimating recombination rates from population genetic data. Genetics **159**(3): 1299-1318.

Feil, E. J. and B. G. Spratt (2001). Recombination and the population structures of bacterial pathogens. Annu. Rev. Microbiol. **55**: 561-590.

Felsenstein, J. (2006). Accuracy of Coalescent Likelihood Estimates: Do We Need More Sites, More Sequences, or More Loci? Mol Biol Evol **23**(3): 691-700.

Gillespie, J. H. (1998). Population Genetics: A Concise Guide. The Johns Hopkins University Press, Baltimore. Ch. 1-2.

Grünwald, N. J., S. B. Goodwin, et al. (2003). Analysis of genotypic diversity data for populations of

microorganisms. Phytopathology **93**(6): 738-746.

Halkett, F., J.-C. Simon, et al. (2005). Tackling the population genetics of clonal and partially clonal organisms. Trends in Ecology & Evolution **20**(4): 194-201.

Huson, D. H. and D. Bryant (2006). Application of phylogenetic networks in evolutionary studies. Molecular Biology and Evolution **23**(2): 254-267.

Jombart, T., S. Devillard, et al. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics **11**: 94.

Hey, J. and R. Nielsen (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics **167**(2): 747-760.

Knowles, L. L. (2009). Statistical phylogeography. Annual Review of Ecology, Evolution and Systematics **40**: 593-612.

Kuhner, M. K. (2009). Coalescent genealogy samplers: windows into population history. Trends in Ecology & Evolution **24**(2): 86-93.

Lemey, Salemi, and Vandamme, eds. (2009) The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing. Cambridge U. Press, Cambridge. Ch. 1, 6. 14.

Li, W.-H. (1997). Molecular Evolution. Sinauer Associates, Inc.

McDonald, B. A. and C. Linde (2002). Pathogen population genetics, evolutionary potential, and durable resistance. Annual Review of Phytopathology **40**: 349-+.

Nordborg, M. (2001). Coalescent theory. Handbook of Statistical Genetics. D. J. Balding, M. J. Bishop and C. Cannings. Chichester, U.K., John Wiley & Sons, Inc.**:** 179–212.

Peever, T. L., S. S. Salimath, et al. (2004). Historical and contemporary multilocus population structure of *Ascochyta rabiei* (teleomorph: *Didymella rabiei*) in the Pacific Northwest of the United States. Molecular Ecology **13**: 291-309.

Pritchard, J. K., M. Stephens, et al. (2000). Inference of population structure using multilocus genotype data. Genetics **155**(2): 945-959.

Schlötterer, C. (2004). The evolution of molecular markers - just a matter of fashion? Nature Reviews Genetics **5**(1): 63-69.

Stukenbrock, E. H., S. Banke, et al. (2007). Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. Molecular Biology and Evolution **24**(2): 398-411.

Wakeley, J. (2005). The Limits of Theoretical Population Genetics. Genetics **169**: 1–7.

Wakeley, J. (2009) Coalescent Theory: An Introduction. Roberts and Company, Greenwood Village, Colorado. Ch 1, 3.