

## PHC 6937 Computational Multi-Omics (1 credit hour)

Summer A: 2021

Delivery Format: Online Synchronous

Course Website: <https://ufl.instructure.com/courses/428973>

---

Instructor Name: Huaizhen Qin, PhD

Room Number: CTRB 4214

Phone Number: 352-294-8545

Email Address: [hqin@ufl.edu](mailto:hqin@ufl.edu)

Office Hours: By appointment

Teaching Assistants: TBD

Preferred Course Communications (e.g. email, office phone): Email

### Prerequisites

PHC 7594 Genetic Epidemiology, PHC 6000 Epidemiology Methods I, PHC 6011 Epidemiology Methods II, and PHC 6050C Biostatistical Methods I. Students must be familiar with R scientific programming languages. This course assumes fundamental competency in genetic epidemiology and statistical genetics vocabulary and principles. Talk to the instructor for a prerequisite waiver or further information.

---

## Purpose and Outcome

### Course Overview

This course will offer master and PhD students with an updated introduction to well-developed and widely used computational algorithms in systems epidemiology. It will introduce the computational approaches to unravel the multi-omics (i.e., genetics, epigenetics, gene expression, and metabolite) mechanisms of complex human diseases. Statistical models and R packages allowing for both unrelated individuals and extended pedigrees in diverse populations will be inspected. This course will cover population and extended pedigree based genome-genome gene mapping, mixed effects Cox regression analyses, and surrogate variable analyses in multi-omics experiments. Public and simulated molecular and phenotypic data sets will be analyzed using the inspected methods for illustration purpose.

### Relation to Program Outcomes

This course is designed for master (MS/MPH) and PhD students who are interested in reliably deciphering the multi-omics basis of complex human diseases. It will update the students with well-developed algorithms and packages which are widely used in systems epidemiology. It will assist the students identify and tackle important challenges in modern systems epidemiology.

### Course Objectives and/or Goals

This course aims to train students to identify, understand and manipulate the well-established analytic methods and R toolkits in computational multi-omics. Upon completion of the course, students will be able to:

- 1) Accurately infer unknown population structures in unrelated individuals from distinct and admixed subpopulations from their DNA sequences
- 2) Visualize explicit (reported) pedigree structure and consistently infer cryptic relatedness between the sampled members from extended pedigrees from their DNA sequences, i.e., multi-generation pedigrees from multiple study centers as can be obtained from multiple dbGap studies
- 3) Accurately separate (unknown) population structure from (cryptic) familial relatedness in the genotypic data matrix of individuals sampled from extended pedigrees
- 4) Consistently estimate the heritability of a complex quantitative trait and conduct reliable whole-genome scan for genetic determinants of the trait after jointly adjusting for the confounding of population structure and familial relatedness
- 5) Conduct reliable whole-genome scans for genetic determinants of binary disease after jointly adjusting for the confounding of case-control imbalance, population structure and familial relatedness
- 6) Control the confounding of relatedness in survival analyses by fitting mixed effects cox models such as the Cox frailty model taking pedigree ID as a random effect and the mixed effects Cox model using genetic relation matrix instead
- 7) Perform surrogate variable analysis to handle unknown sources of variation in high-throughput experiments, e.g., DNA methylation, RNA sequence, brain imaging
- 8) Eliminate the bias and inflation due to latent confounders in epigenome-and transcriptome-wide association studies

### **Instructional Methods**

The instructor will lecture on methodological details and illustrate the applications of computational packages. Working together with the instructor in class and lab sessions, students will learn how to use the algorithms and packages to dig multi-omics data and interpret the results. Moreover, students will give presentations on selected algorithms and packages for modern systems epidemiology.

#### *What is expected of you?*

You are expected to actively engage in the course throughout the semester. You must come to each class/lab prepared by completing all out-of-class assignments. This preparation gives you the knowledge or practice needed to engage in higher levels of learning during the live class sessions. If you are not prepared for the face-to-face sessions, you may struggle to keep pace with the activities occurring in the live sessions, and it is unlikely that you will reach the higher learning goals of the course. Similarly, you are expected to actively participate in the live classes. Your participation fosters a rich course experience for you and your peers that facilitates overall mastery of the course objectives.

---

## Description of Course Content

### Topical Outline/Course Schedule

Week	Date(s)	Topic(s)	Readings	Assignments
1	5/11/21 Period 5 2:00p – 3:15p	<i>Basic Concepts</i> : genomic markers, alleles, linkage disequilibrium, phenotypes, association mapping, covariates and confounders, latent variables	Article [1]	
	5/11/21 Period 6 3:30p – 4:45p	Population Structure (PS): distinct subpopulations, admixed minorities, revealing population structure using R package <i>mixOmics</i> , <i>PC-AiR</i> , or the <i>PLINK</i> & <i>GCTA</i>	Articles [2-6]	
2	5/18/21 Period 5 2:00p – 3:15p	<i>Familial Relatedness</i> : visualize explicit (reported) pedigree structure and detect cryptic relatedness to correct pedigree plots	Articles [7-9]	<b>Assignment 1:</b> Compare kinship2 and PLINK ped
	5/18/21 Period 6 3:30p – 4:45p	<i>R package kinship2: explicit extensive pedigree; GCTA tool: genotypic relation matrix (GRM); REAP: cryptic relatedness among admixed subjects; PC-Relate</i>		
3	5/25/21 Period 5 2:00p – 3:15p	<i>Whole-genome Gene Mapping</i> : heritability of a <i>quantitative trait</i> , calibrate case-control imbalance, population structure, and familial relatedness	Articles [10, 11]	<b>Assignment 2:</b> Compare SMMAT and SAIGE
	5/25/21 Period 6 3:30p – 4:45p	R packages <i>SMMAT</i> & <i>SAIGE</i> (calibrate multiple mixed effects in gene mapping for quantitative traits and binary diseases)		
4	6/1/21 Period 5 2:00p – 3:15p	Mixed Effects Cox Models: Cox frailty model taking pedigree ID as a random effect, mixed effects Cox model using genetic relation matrix instead	Articles [12, 13],	<b>Assignment 3:</b> Compare survival frailty and coxme
	6/1/21 Period 6 3:30p – 4:45p	R package <i>survival</i> : Cox frailty regression using pedigree ID as a random effect; R package <i>coxme</i> : mixed effects Cox regression using GRM		
5	6/8/21 Period 5 2:00p – 3:15p	<i>Surrogate Variable Analysis</i> : handle unknown sources of variation in high-throughput experiments, e.g., DNA methylation, RNA sequence, brain imaging	Articles [14, 15]	<b>Student presentations</b>
	6/8/21 Period 6 3:30p – 4:45p	<i>R package sva</i> : construct surrogate variable and remove the effects of unwanted variation in high-throughput experiments		
6	6/15/21 Period 5 2:00p – 3:15p	Controlling bias and inflation due to latent confounders in epigenome-and transcriptome-wide association studies	Articles [16, 17]	<b>Student presentations</b>
	6/15/21 Period 6 3:30p – 4:45p	R package <i>BACON</i> : controlling bias and inflation using empirical null distribution; R package <i>LFMM2</i> : fitting latent factor mixed-effects models for confounder adjustment in genome and epigenome-wide association studies		

### Course Materials and Technology

#### Readings for Discussions and Presentations

- Hirschhorn, J.N. and M.J. Daly, *Genome-wide association studies for common diseases and complex traits*. Nature reviews genetics, 2005. 6(2): p. 95-108.
- Rohart, F., et al., *mixOmics: An R package for 'omics feature selection and multiple data integration*. PLoS computational biology, 2017. 13(11): p. e1005752.

3. Purcell, S., et al., *PLINK: a tool set for whole-genome association and population-based linkage analyses*. The American Journal of Human Genetics, 2007. **81**(3): p. 559-575.
4. Price, A.L., et al., *Principal components analysis corrects for stratification in genome-wide association studies*. Nature genetics, 2006. **38**(8): p. 904-909.
5. Yang, J., et al., *GCTA: a tool for genome-wide complex trait analysis*. The American Journal of Human Genetics, 2011. **88**(1): p. 76-82.
6. Conomos, M.P., M.B. Miller, and T.A. Thornton, *Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness*. Genetic epidemiology, 2015. **39**(4): p. 276-293.
7. Astle, W. and D.J. Balding, *Population structure and cryptic relatedness in genetic association studies*. Statistical Science, 2009. **24**(4): p. 451-471.
8. Sinnwell, J.P., T.M. Therneau, and D.J. Schaid, *The kinship2 R package for pedigree data*. Human heredity, 2014. **78**(2): p. 91-93.
9. Conomos, M.P., et al., *Model-free estimation of recent genetic relatedness*. The American Journal of Human Genetics, 2016. **98**(1): p. 127-148.
10. Chen, H., et al., *Efficient variant set mixed model association tests for continuous and binary traits in large-scale whole-genome sequencing studies*. The American Journal of Human Genetics, 2019. **104**(2): p. 260-274.
11. Zhou, W., et al., *Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies*. Nature genetics, 2018. **50**(9): p. 1335.
12. Therneau, T., *Mixed effects Cox models*. CRAN repository, 2019.
13. Pankratz, V.S., M. De Andrade, and T.M. Therneau, *Random-effects Cox proportional hazards model: General variance components methods for time-to-event data*. Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society, 2005. **28**(2): p. 97-109.
14. Leek, J.T., et al., *The sva package for removing batch effects and other unwanted variation in high-throughput experiments*. Bioinformatics, 2012. **28**(6): p. 882-883.
15. Leek, J.T., et al., *sva: Surrogate Variable Analysis R package version 3.10.0*. DOI, 2014. **10**: p. B9.
16. van Iterson, M., E.W. van Zwet, and B.T. Heijmans, *Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution*. Genome biology, 2017. **18**(1): p. 1-13.
17. Caye, K. and O. François, *LFMM 2.0: Latent factor models for confounder adjustment in genome and epigenome-wide association studies*. Biorxiv, 2018: p. 255893.

### **Data to be Used and IRB**

For this course, only simulated data will be used for illustration and examination purposes and thus IRB approval is not required. However, you will have to apply for an IRB approval if you wish to analyze de-identified public human data by using methods or pipelines introduced in the class. For conducting research on human subjects, you can apply for IRB approval through the website (<https://my.irb.u.edu>). I wish that everyone can make important scientific discoveries.

### **Computer Technology**

Students are expected to bring a laptop with R and other software packages. R is free and more flexible for data analysis and visualization. The course work can be successfully completed R. But students are welcome to use whichever pipeline.

For technical support for this class, not related to R programming, please contact the UF Help Desk at:

- [Learning-support@ufl.edu](mailto:Learning-support@ufl.edu)
- (352) 392-HELP - select option 2
- <https://lss.at.ufl.edu/help.shtml>

## Academic Requirements and Grading

### Assignments

Individual evaluation will be based on performance in class participation, homework assignments, one oral presentation, and two exams (a midterm and a final).

- Homework (60%): Students are required to independently complete and turn in answers for all assignments on the due dates listed in **Grading** below. No overdue answers will be accepted or graded, except in emergency situations with proof. Homework will be graded on the basis of subject knowledge, analytics, innovation and organization.
- Oral presentation (40%): Students are required to intensively read and orally present the original research articles group-wise or individually. Each presentation will be 30 minutes, including question and comment time. Oral presentations will be graded on the basis of <https://people.clas.ufl.edu/glord/files/Oral-presentation-grading-rubric.pdf>

### Grading

Requirement	Due date	% of final grade
Homework		60% total
Assignment 1	5/25/21	20%
Assignment 2	6/1/21	20%
Assignment 3	6/8/21	20%
Oral presentation	6/8/21 & 6/15/21	40%

Point system used (i.e., how do course points translate into letter grades).

Points Earned	Letter Grade
93-100	A
90-92	A-
87-89	B+
83-86	B
80-82	B-
77-79	C+
73-76	C
70-72	C-
67-69	D+
63-66	D
60-62	D-
Below 60	E

Please be aware that a C- is not an acceptable grade for graduate students. The GPA for graduate students must be 3.0 based on 5000 level courses and above to graduate. A grade of C counts toward a graduate degree only if based on credits in courses numbered 5000 or higher that have been earned with a B+ or higher.

Letter Grade	Grade Points
A	4.0
A-	3.67
B+	3.33
B	3.0
B-	2.67
C+	2.33
C	2.0
C-	1.67
D+	1.33
D	1.0
D-	0.67
E	0.0
WF	0.0
I	0.0
NG	0.0
S-U	0.0

More information on UF grading policy may be found at:

<http://gradcatalog.ufl.edu/content.php?catoid=10&navoid=2020#grades>

### **Exam Policy**

Both midterm and final exams will be conducted in-class and open-book. The exams will be in the format of 60 percent multiple choice and 40 percent analysis-based long answer. All topics covered in class sessions will be possibly examined. Laptop use is permitted but cell phone use is prohibited during the examination time. Students who leave exam room for any reason during an exam must put all exam materials face-down and close laptops, and are not permitted to take any examination materials outside the room. Students must independently solve exam problems and are prohibited to seek any assistance from any others in any way. Instructors/proctors will be present in exam room periodically during each exam to evade all exam security measures.

### **Policy Related to Make up Exams or Other Work**

Makeup work will be allowed with instructor's permission on an individual basis after an excused absence. Excused absences must be consistent with university policies in the Graduate Catalog (<http://gradcatalog.ufl.edu/content.php?catoid=10&navoid=2020#attendance>). Additional information can be found here: <https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx>

Please note: Any requests for make-ups due to technical issues MUST be accompanied by the UF Computing help desk (<http://helpdesk.ufl.edu/>) correspondence. You MUST e-mail me within 24 hours of the technical difficulty if you wish to request a make-up.

### **Policy Related to Required Class Attendance**

Students are expected to attend all class sessions and actively participate in all discussions. Students who cannot attend a class meeting should inform the instructor via email prior to the date of the class, or on the day of the absence for illness or emergency. Being more than 20 minutes late for a meeting will be taken as an absence.

Please note all faculty are bound by the UF policy for excused absences. Excused absences must be consistent with university policies in the Graduate Catalog

(<http://gradcatalog.ufl.edu/content.php?catoid=10&navoid=2020#attendance>). Additional information can be found here: <https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx>

---

## Student Expectations, Roles, and Opportunities for Input

### Expectations Regarding Course Behavior

Please come to class on time and be prepared to stay until the time scheduled as the end of class. Your investment in the degree is worth maximizing your in-class experience, and we expect to provide materials that utilize the full, scheduled class times. The use of cell phones is not permitted. Please turn them off or, if you expect urgent calls, set them to vibrate.

### Communication Guidelines

The best way to communicate with the instructor is by email, either through the Canvas site or through the UF email system. Please give the instructor up to 24 hours during the work week to respond to your email. Email correspondence should follow the etiquette of business emails (see UF's *Netiquette Guide for Online Courses* for guidance at <http://teach.ufl.edu/wp-content/uploads/2012/08/NetiquetteGuideforOnlineCourses.pdf>). If you would like to meet in person outside of office hours then please make an appointment to see the instructor; this ensures that you will have an uninterrupted time to meet.

### Academic Integrity

Students are expected to act in accordance with the University of Florida policy on academic integrity. As a student at the University of Florida, you have committed yourself to uphold the Honor Code, which includes the following 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.”**

You are expected to exhibit behavior consistent with this commitment to the UF academic community, and on all work submitted for credit at the University of Florida, the following pledge is either required or implied:

**“On my honor, I have neither given nor received unauthorized aid in doing this assignment.”**

It is your individual responsibility to know and comply with all university policies and procedures regarding academic integrity and the Student Honor Code. Violations of the Honor Code at the University of Florida will not be tolerated. Violations will be reported to the Dean of Students Office for consideration of disciplinary action. For additional information regarding Academic Integrity, please see Student Conduct and Honor Code or the Graduate Student Website for additional details:

<https://www.dso.ufl.edu/sccr/process/student-conduct-honor-code/>  
<http://gradschool.ufl.edu/students/introduction.html>

Please remember cheating, lying, misrepresentation, or plagiarism in any form is unacceptable and inexcusable behavior.

### Citations and Plagiarism

The two key purposes of citation are to: 1) give appropriate credit to the authors of information, research findings, and/or ideas (and avoid plagiarism), and 2) facilitate access by your readers to the sources you use in your research.

Quotations: When directly quoting an outside source, the borrowed text, regardless of the amount, must be surrounded by quotation marks or block quoted. Quoted text over two lines in length should be single-spaced and indented beyond the normal margins. Every quote must include a source—the author, title, volume, page numbers, etc.—whether an internal reference, footnote, or endnote is used in conjunction with a bibliography page.

Paraphrasing or Citing an Idea: When summarizing an outside source in your own words or citing another person's ideas, quotation marks are not necessary, but the source must be included. This includes, but is not confined to, personal communications from other students, faculty members, experts in the field, summarized ideas from published or unpublished resource, and primary methods derived from published or unpublished sources. Use the general concept of “when in doubt – cite.”

Plagiarism is a serious violation of the academic honesty policy of the University. If a student plagiarizes others' material or ideas, UF Policies on Honesty and honor code violations, noted above, will be followed.

Generally speaking, the three keys of acceptable citation practice are: 1) thoroughness, 2) accuracy, and 3) consistency. In other words, be sure to fully cite all sources used (thoroughness), be accurate in the citation information provided, and be consistent in the citation style you adopt. All references should include the following elements: 1) last names along with first and middle initials; 2) full title of reference; 3) name of journal or book; 4) publication city, publisher, volume, and date; and 5) page numbers referenced. When citing information from the Internet, include the WWW address at the end, with the “access date” (i.e., when you obtained the information), just as you would list the document number and date for all public documents. When citing ideas or words from an individual that are not published, you can write “personal communication” along with the person's name and date of communication.

### **Use of Unauthorized Assistance Resources**

As a graduate student at UF, you are expected to present your own work for grading. Unauthorized sources of help, including commercially available software and services, are not allowed. Even though the students will not be graded on their grammar, it is expected that as graduate students you will have sufficient English language skills to convey your thoughts in organized and understandable manner. If the assignment is unreadable, it will not be graded and will be assigned zero points. Use of unauthorized assistance resources will result in zero points on the written assignment and a report to the Dean of Student's Office. If English is your second language, you may visit the UF Writing Program Website to learn about available help.

### **Online Faculty Course Evaluation Process**

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at <https://gatorevals.aa.ufl.edu/students/>. Students will be notified when the evaluation period opens, and can complete evaluations through the email they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via <https://ufl.bluera.com/ufl/>. Summaries of course evaluation results are available to students at <https://gatorevals.aa.ufl.edu/public-results/>.

### **Online Synchronous Sessions:**

Our class sessions may be audio visually recorded for students in the class to refer back and for enrolled students who are unable to attend live. Students who participate with their camera engaged or utilize a profile image are agreeing to have their video or image recorded. If you are unwilling to consent to have your profile or video image recorded, be sure to keep your camera off and do not use a profile image. Likewise, students who un-mute during class and participate orally are agreeing to have their voices recorded. If you are not willing to consent to have your voice recorded during class, you will need to keep your mute button activated and communicate exclusively using the "chat" feature, which allows students to type questions and comments



live. The chat will not be recorded or shared. As in all courses, unauthorized recording and unauthorized sharing of recorded materials is prohibited.

### **Policy Related to Guests Attending Class**

Only registered students are permitted to attend class. However, we recognize that students who are caretakers may face occasional unexpected challenges creating attendance barriers. Therefore, by exception, a department chair or his or her designee (e.g., instructors) may grant a student permission to bring a guest(s) for a total of two class sessions per semester. This is two sessions total across all courses. No further extensions will be granted. Please note that guests are **not** permitted to attend either cadaver or wet labs. Students are responsible for course material regardless of attendance. For additional information, please review the Classroom Guests of Students policy in its entirety. Link to full policy:

<http://facstaff.phhp.ufl.edu/services/resourceguide/getstarted.htm>

---

## **SUPPORT SERVICES**

### **Accommodations for Students with Disabilities**

If you require classroom accommodation because of a disability, it is strongly recommended you register with the Dean of Students Office <http://www.dso.ufl.edu> within the first week of class or as soon as you believe you might be eligible for accommodations. The Dean of Students Office will provide documentation of accommodations to you, which you must then give to me as the instructor of the course to receive accommodations. Please do this as soon as possible after you receive the letter. Students with disabilities should follow this procedure as early as possible in the semester. The College is committed to providing reasonable accommodations to assist students in their coursework.

### **Counseling and Student Health**

Students sometimes experience stress from academic expectations and/or personal and interpersonal issues that may interfere with their academic performance. If you find yourself facing issues that have the potential to or are already negatively affecting your coursework, you are encouraged to talk with an instructor and/or seek help through University resources available to you.

- The Counseling and Wellness Center 352-392-1575 offers a variety of support services such as psychological assessment and intervention and assistance for math and test anxiety. Visit their web site for more information: <http://www.counseling.ufl.edu>. On line and in person assistance is available.
- You Matter We Care website: <http://www.umatter.ufl.edu/>. If you are feeling overwhelmed or stressed, you can reach out for help through the You Matter We Care website, which is staffed by Dean of Students and Counseling Center personnel.
- The Student Health Care Center at Shands is a satellite clinic of the main Student Health Care Center located on Fletcher Drive on campus. Student Health at Shands offers a variety of clinical services. The clinic is located on the second floor of the Dental Tower in the Health Science Center. For more information, contact the clinic at 392-0627 or check out the web site at: <https://shcc.ufl.edu/>
- Crisis intervention is always available 24/7 from:  
Alachua County Crisis Center:  
(352) 264-6789  
<http://www.alachuacounty.us/DEPTS/CSS/CRISISCENTER/Pages/CrisisCenter.aspx>

Do not wait until you reach a crisis to come in and talk with us. We have helped many students through stressful situations impacting their academic performance. You are not alone so do not be afraid to ask for assistance.

### **Inclusive Learning Environment**

Public health and health professions are based on the belief in human dignity and on respect for the individual. As we share our personal beliefs inside or outside of the classroom, it is always with the understanding that we value and respect diversity of background, experience, and opinion, where every individual feels valued. We believe in, and promote, openness and tolerance of differences in ethnicity and culture, and we respect differing personal, spiritual, religious and political values. We further believe that celebrating such diversity enriches the quality of the educational experiences we provide our students and enhances our own personal and professional relationships. We embrace The University of Florida's Non-Discrimination Policy, which reads, "The University shall actively promote equal opportunity policies and practices conforming to laws against discrimination. The University is committed to non-discrimination with respect to race, creed, color, religion, age, disability, sex, sexual orientation, gender identity and expression, marital status, national origin, political opinions or affiliations, genetic information and veteran status as protected under the Vietnam Era Veterans' Readjustment Assistance Act." If you have questions or concerns about your rights and responsibilities for inclusive learning environment, please see your instructor or refer to the Office of Multicultural & Diversity Affairs website: [www.multicultural.ufl.edu](http://www.multicultural.ufl.edu)

---