

H3 Africa Program of the National Human Genome Research Institute,
NIH, USA, and Wellcome Trust, United Kingdom

2015 Proteomics Training Course

J Craig Venter Institute, Campus in Rockville, Maryland,
October 4 - 9, 2015

Preliminary Course Program

National Human Genome Research Institute (NHGRI)
J. Craig Venter Institute (JCVI)

J. Craig Venter[®]
I N S T I T U T E

J. Craig Venter Institute (JCVI)

The J. Craig Venter Institute was formed in October 2006 through the merger of several affiliated and legacy organizations — The Institute for Genomic Research (TIGR) and The Center for the Advancement of Genomics (TCAG), The J. Craig Venter Science Foundation, The Joint Technology Center, and the Institute for Biological Energy Alternatives (IBEA). Today all these organizations have become one large multidisciplinary genomic-focused organization. With approximately 200 scientists and staff, more than 250,000 square feet of laboratory space, and locations in Rockville, Maryland and San Diego, California, the new JCVI is a world leader in genomic research.

Proteomics

Proteomics is a research discipline increasingly used in a variety of fields in biology driven by powerful technology developments in mass spectrometry and computational analysis of mass spectral data and supported by the sequence analysis and annotation of genomes. Examples of scientific disciplines where proteomics is useful are the characterization of cellular function and architecture, the determination of interactions between host and pathogen, the identification of biomarkers for non-communicable diseases (diagnosis, prognosis, prediction, drug treatment efficacy, drug toxicity), and the simple identification of proteins and their modifications.

Website: www.jcvi.org

Main workshop organizers:

Rembert Pieper, Ph.D. (Associate Professor, J. Craig Venter Institute)

<http://www.jcvi.org/cms/about/bios/rpieper>

rpieper@jcvi.org

Yanbao Yu, Ph.D. (Staff Scientist, J. Craig Venter Institute)

<http://www.jcvi.org/cms/about/bios/yayu/>

yayu@jcvi.org

Workshop co-organizers:

Keehwan Kwon, Ph.D. (Assistant Professor, J. Craig Venter Institute)

<http://www.jcvi.org/cms/about/bios/kkwon/>

kkwon@jcvi.org

Rajagopala Venkatappa, Ph.D. (Assistant Professor, J. Craig Venter Institute)

<http://www.jcvi.org/cms/about/bios/rvenkatappa>

rajagopala@jcvi.org

Tim Stayeas

Network Team Lead, Information Technology, J. Craig Venter Institute

tstayeas@jcvi.org

Logistic information:

Inquiries

Elaine Fox

Office: +1.301.795.7401

Email: efox@jcv.org

Venue

J. Craig Venter Institute, 9704 Medical Center Drive, Rockville, MD 20850

Dates

October 4 to October 9, 2015

Please note that the Rockville campus of the JCVI is located 19 miles from the location of the H3Africa conference taking place from October 11 to October 15, 2015 (Omni Shoreham Hotel, 2500 Calvert St NW, Washington, DC 20008). There are two hotels across the street from JCVI so that public transportation is not needed during the time of the workshop provided that a participant will stay in one of those hotels. JCVI will help with the reservation of accommodations in one of these hotels.

Meals

The following meals will be provided for workshop participants.

Sunday	October 4, 2015	Lunch & Breaks
Monday	October 5, 2015	Lunch & Breaks
Tuesday	October 6, 2015	Lunch & Breaks
Wednesday	October 7, 2015	Lunch & Breaks
Thursday	October 8, 2015	Lunch & Breaks
Friday	October 9, 2015	Lunch & Breaks

Information for participants:

The question arises as to whether the proteomics course is relevant to specific projects of the H3Africa program. The training program can be customized to a certain degree if the organizers have sufficient time to prepare for that. The organizers can prepare and adapt seminars useful in the context of the participants' projects and interests to some extent. The organizers can analyze samples during the course that are of high interest to the participants as long as the participants make provisions for permitted and safe import of their own samples. All participants will have individual computer workstations for the bioinformatics training components. All participants will have the opportunity to learn one of the dominant proteomic techniques hands-on in the laboratory, shotgun proteomics with mass spectrometry as the protein identification technology.

Brief training course summary:

The proteomics training program focusses on shotgun proteomics and mass spectrometry. An experiment involves a set of experimental and computational steps for which the minimal time is 5 days. The experimental work begins with the generation of a lysate or concentrate of proteins that is quantified prior to proceeding to the filter aided sample preparation (FASP) stage. FASP generates the tryptic peptide mixtures which are ideal for shotgun proteomic analysis on a nano-flow liquid chromatography nano-spray mass spectrometry (nLC-MS/MS) workstation. The next stage is the nLC-MS/MS experiment which is followed by a database search to match the mass-over-charge (m/z) ratios observed by the MS to tryptic peptides predicted from the database that contains the protein sequence information relevant to the source of the sample. The database search returns information of the proteins present, their estimated amounts and the confidence with which they are identified. This information is obtained by a software tool. We will introduce two software tools that provide more accurate information of protein quantities (MaxQuant) and allow statistical analyses of proteomics data (MeV). At the end of the workshop, we will review the data participants obtained and discuss applications to their projects and areas of biological interest.

Analysis Stages in Quantitative Proteomics

Urine sample collection

Isolation of
- Pellet fraction
- Soluble fraction

Protein denaturation → **Protein digestion**

Peptide separation

Electrospray

MS/MS

Data analysis with search engine

```

1 HARTTPIARK RHIGISAMID AGRTTTERI LFTYOVNKI QEVHDQAIR
51 DHEEZEQER ITITSIATTA FSGRMKQYE RSHNIISTP GRVFTTKEV
101 RSHFVLDGAV MYVCANVGQV PQSETVWQA NKYKVPRIAF VNHSDRIGAN
151 FLKVVHQIKT DLGAMPVPLQ LAIGAEHFT GVVDLVKKA INNNDADQV
201 TFEYDIPAD NVELANWHD NLESAAEAS EELNRYLGG ZELTEADIQ
251 ALKQVLMRE IILVTCGAF YKQQAHLG HWIDYLSFV DIFALNIGLD
301 DGEDTAEEM ASDDEFFSAL AFKIATDPFV GHLTFKVVYS GVNSGDTVL
351 NSVKAAREF GRIVQNHAK REEIKVVRG DIAAALGLD VTTGTLCDP
401 DAFILLEHE FPEPVISLAV EPRKADQEK NGLLGLAK EDPSFVWTD
451 RESQITLIG NGLDLIDIV DRKRFNWE ARVQSPQAV SETIQRYVD
501 VEGHAKQEG GRQYGRVVI DNYPLEPSN PKGYEFINDI KGGVIGEVI
551 PAVDKIQEQ LKAGPLAGV VVDNIGLHF GYHVDVDSSE LAFKLAASIA
601 FREGFRKAKP VLLEPDKVE VETPEENTG VIQGLSRRG RLKQGESEVT
651 QVKHMANPL SDFQVATQL RSLTQKAST TREFLAYEA PSNVAQVIE
                    
```

Relative quantitative analysis

- Annexin A1 rel. quantity: 2345
- Alkyl hydroperoxide reductase subunit C rel. quantity: 12309

J. Craig Venter
INSTITUTE

Shotgun proteomic analyses stages. The sample source could be different from a human urine sample, and may require different methods in the sample preparation and database searches stages.

Preliminary proteomics training course program

Sunday October 4, 2015

- 12:00 noon Arrival and Registration
- 1:00 pm Lunch, Welcome and Opening Remarks
- 1:30 pm Tour of the J. Craig Venter Institute and the proteomics and bioinformatics laboratory
- 2:30 pm Introduction into proteomics (lecture/seminar)
- 3:30 pm Description of the proteomics course components
- 4:00 pm Preparation for the first day in the laboratory
- 5:00 pm Discussion of participant-specific projects and proteomic analyses (if applicable)
- 6:00 pm Joint dinner or reception, place to be determined

Monday October 5, 2015

- 8:30 am Proteomics laboratory: preparation of protein extracts and concentrates (this module may include cell or tissue lysis experiments, body fluid concentration, and protein extract analysis by protein concentration measurements and SDS-PAGE)
- 12:00 noon Lunch
- 1:00 pm Proteomics laboratory: preparation of protein extracts and concentrates (continued)
- 3:00 pm Discussion of participants' results
- 4:00 pm Discussion of a journal article focused on proteomic sample preparation
- 6:00 pm Activities adjourned

Tuesday October 6, 2015

- 8:30 am Proteomics laboratory: FASP filter-aided sample preparation (this module includes the enzymatic generation of peptide mixtures later analyzed by shotgun proteomics experiments)
- 12:00 noon Lunch
- 1:00 pm Proteomics laboratory: FASP filter-aided sample preparation (continued)
- 3:00 pm Seminar on sample preparation for proteomics (expected results, troubleshooting, sample-specific differences, experiment optimization, alternative methods)

4:30 pm Discussion of the participants' results

6:00 pm Activities adjourned

Wednesday October 7, 2015

8:30 am Proteomics laboratory: filter-aided sample preparation (continued)

9:30 am Seminar on principles of liquid chromatography mass spectrometry

11:00 am Proteomics laboratory: mass spectrometry for shotgun proteomics (laboratory practice, Q-Exactive mass spectrometer; setting parameters for data-dependent mass analysis)

12:00 noon Lunch

1:00 pm Proteomics laboratory: mass spectrometry for shotgun proteomics applications (practical application using test samples)

3:00 pm Proteomics laboratory: Stage-Tips (micro-analytical clean-up process for peptide mixtures)

5:00 pm Discussion of the participants' results

6:00 pm Activities adjourned

Thursday October 8, 2015

8:30 am Proteomics laboratory: shotgun proteomics (using participants' samples; practice of LC-MS/MS experiments and database searches)

12:00 noon Lunch

1:00 pm Proteomics laboratory: Computational analysis of test sample proteomes (discussion of results of the database searches: Proteome Discoverer software)

3:00 pm Seminar on shotgun proteomics (expected results, troubleshooting, sample-specific differences, experiment optimization, alternative methods)

5:00 pm Discussion of the results from test samples

6:00 pm Activities adjourned

Friday October 9, 2015

8:30 am Proteomics laboratory: Computational analysis of proteomes from participant samples (explain results of database searches in Proteome Discoverer software; discussion of high vs. low resolution MS and computational methods to assess data quality statistically)

10:30 am	<u>Bioinformatics laboratory:</u> Data analysis on individual basis (each participant is mentored to examine data quality, level of information on the analyzed proteome and quantification of proteome)
12:00 noon	Lunch
1:00 pm	<u>Bioinformatics laboratory:</u> Continued data analysis using Thermo Discoverer, MeV and Max Quant software tools (statistical and biological analyses of participants' proteomic data)
4:00 pm	Initiation of a discussion of ideas how participants can use shotgun proteomics in the context of their research priorities
6:00 pm	Activities adjourned