

Course overview

Bioinformatic Analysis for Cancer Genomics

15/12/2024

Phuc Loi Luu, PhD

Email: Luu.p.loi@gmail.com

Zalo: 0901802182

Content

- The Teams help US!
- What we will learn from the course?
- What you can do after the course?
- Housekeeping: Github and youtube

TEACHING ASSISTANT TEAM (TRỢ GIẢNG)



Nguyễn Quang Khải
nguyenkhai12330@gmail.com



Hà Gia Huy
hagiahuy311@gmail.com



Hồ Phú Quý
hophuquy0944@gmail.com



BS. Hoàng Sơn
hoangsonlvt710@gmail.com



BS. Minh Nguyễn
ducminhnguyenle30091996@gmail.com



ThS. Nguyễn Quỳnh Như
Lucianhu@gmail.com



ThS. Trần Thị Mỹ Qui
ttmqui1912@gmail.com



Duy Đào
khuongduying@gmail.com



ThS. Ngô Đại Phú
phudaingobio@gmail.com



Nguyễn Mạnh Hùng
hungtrangan141@gmail.com



Lê Văn Giàu
lgiau6366@gmail.com



Phạm Văn Quan
quanpham0512@gmail.com



Hoàng Kim
vanhohoangkim@gmail.com



ThS. BS. Đào Ngọc Bắc
daongocbac2020@gmail.com



TS. Trịnh Vạn Ngử
VNTrinh@mdanderson.org

LOGISTIC TEAM (HẬU CẦN)



Hoàng Kim
vanhohoangkim@gmail.com

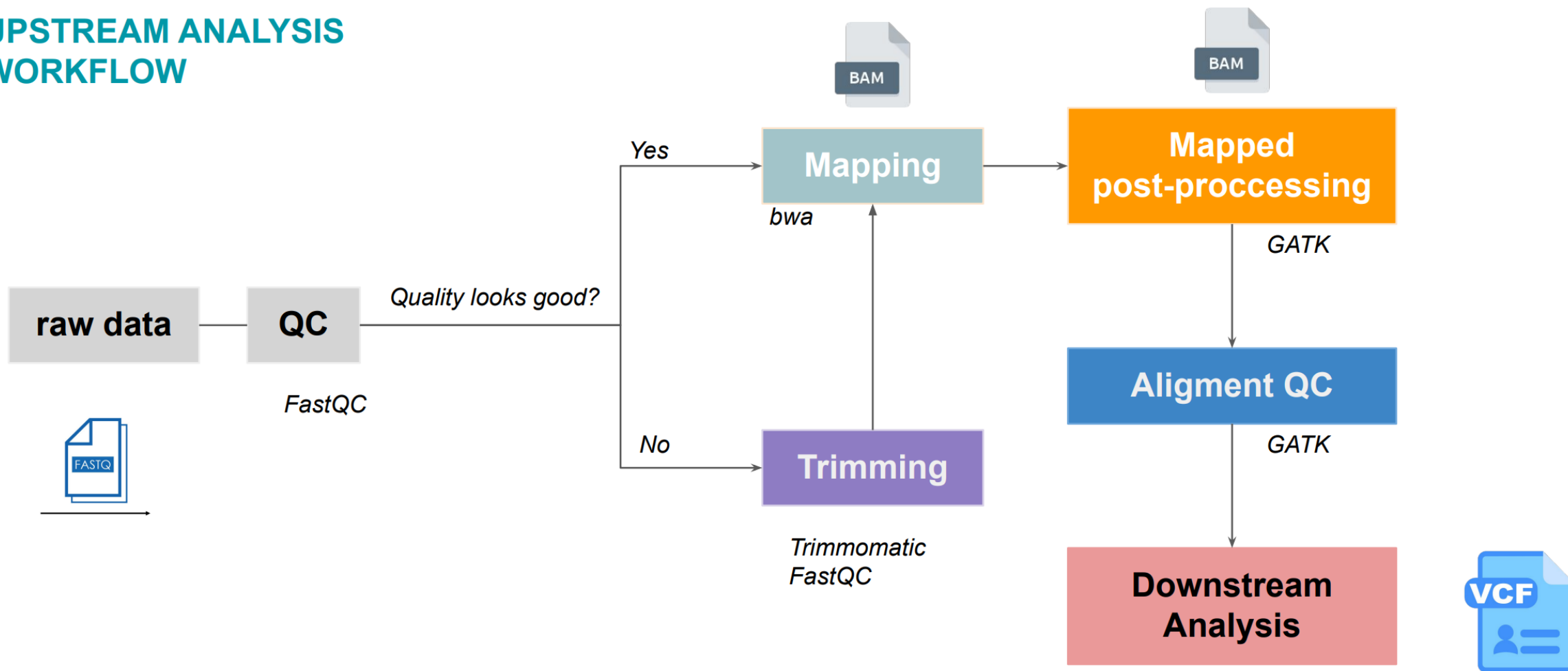


Nguyễn Lê Quang
quanglequang2003@gmail.com

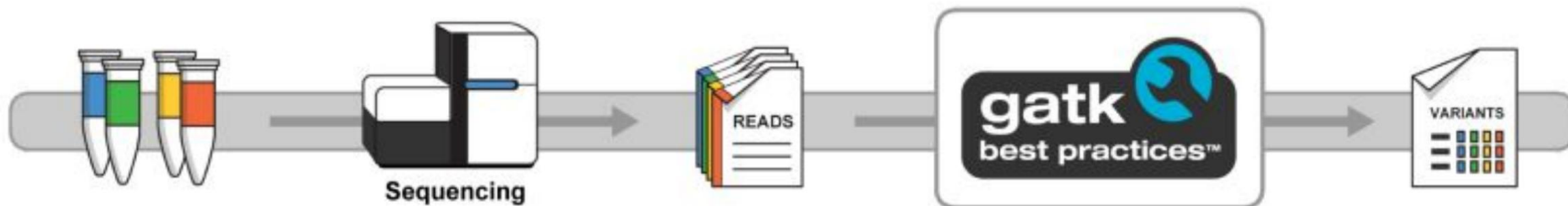


Lê Giàu
lgiau6366@gmail.com

UPSTREAM ANALYSIS WORKFLOW



Workflows for all major variant classes



	GERMLINE	SOMATIC
SNPs & INDELS	HaplotypeCaller GVCF	Mutect2
Copy Number	GATK gCNV	GATK CNV + aCNV
Structure Variation	GATK SVDDiscovery (beta)	(planned)

Methods

No	Sequencing Platforms	Read mapping	Single Nucleotide Variant (SNV) Calling	Structural Variant (SV) and Copy Number Variant (CNV) Calling	Variant Annotation
1	Illumina	BWA-mem, Bowtie2, SOAP2, Minimap2, BLASR, NGMLR ...	SAMtools, Deepvariant, GATK, SpeedSeq ...	DELLY, BreakDancer, LUMPY, GRIDSS, TARDIS, SURVIVOR, PBHoney, NanoSV, Manta, CNVnator ...	HPO, ClinVar, Varsome, Gnomad, Cosmic, ANNOVAR ...
2	Onso (Short read, PacBio) Revio (Long read, PacBio)				
3	Ion Torrent PGM				
4	MGI Cyclone-SEQ (Long read)				
5	GeneMind				
6	Aviti (Element Biosciences)				
7	ONT (Long read)				

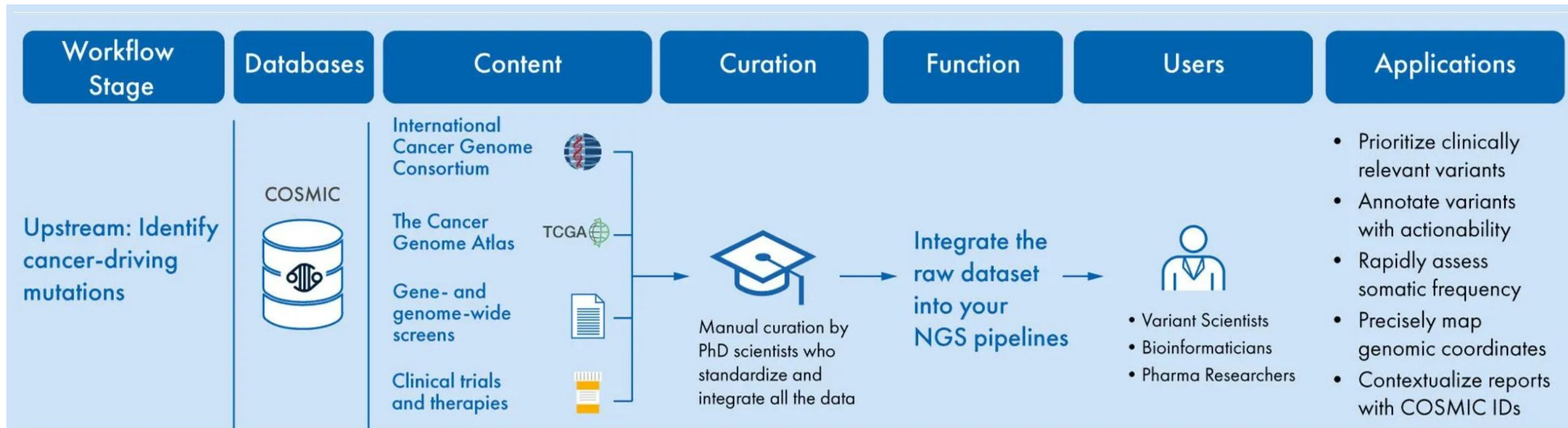
What we will learn from the course?

1. Install and use Ubuntu or Google Colab
2. Hands-on programming with Bash and R, Linux command lines, and some bioinformatic tools
3. Manipulate Next Generation Sequencing data of DNA in both research and clinical application
4. Plot DNA-seq data with ggplots and do statistically analyzing with R
5. Downstream analysis of DNA-seq
6. Read publications of DNA-seq research
7. Discuss the DNA-seq testing reports
8. Search for driver and passenger mutations in cancers

What you can do after the course?

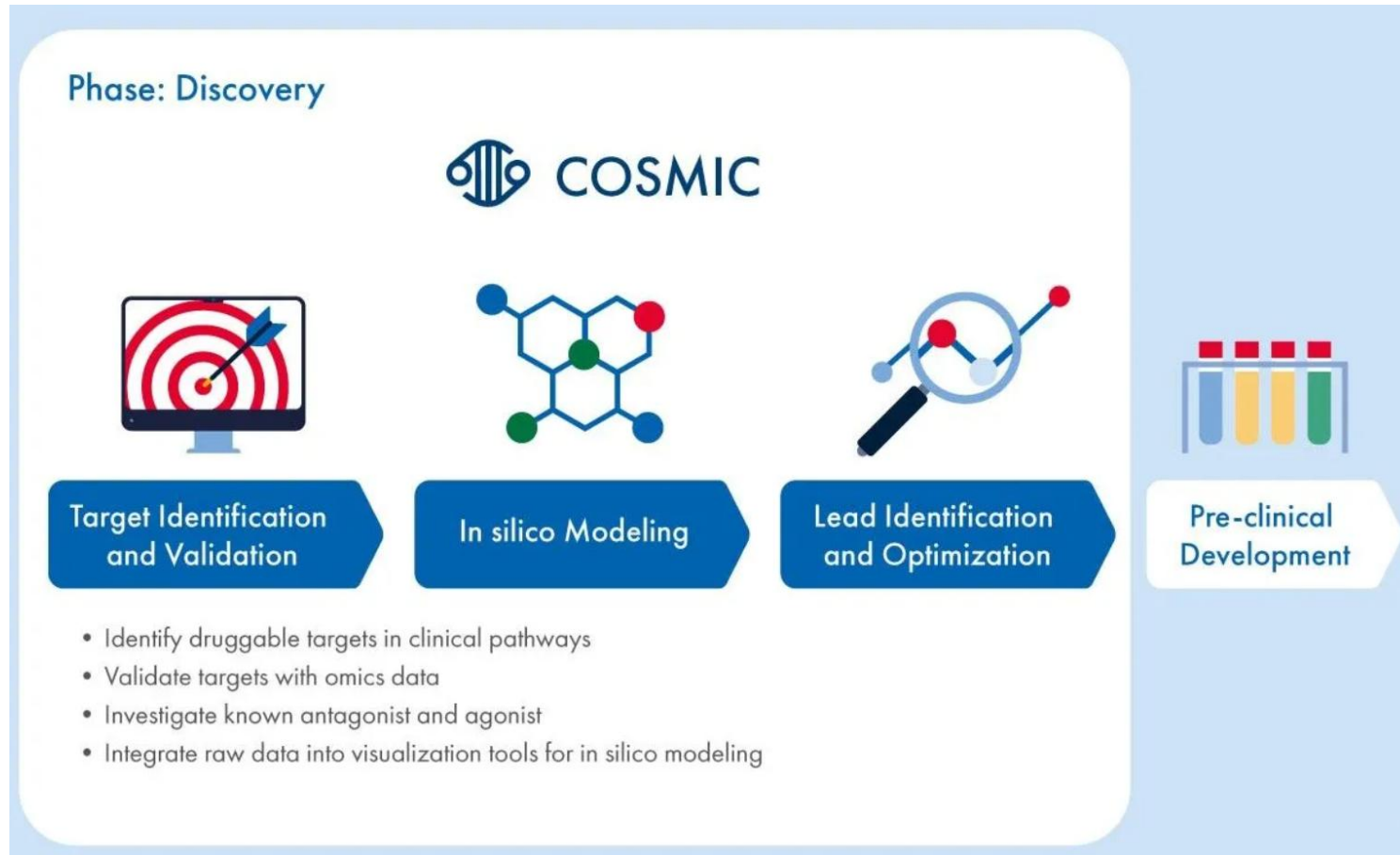
1. Use Ubuntu or Google Colab
2. Master programming language (Bash and R), Linux command line, and some bioinformatic tools
3. Manipulate Next Generation Sequencing data of DNA for both research and clinical application
4. Plot ready-for-publication figures and do statistically analyzing for a manuscript
5. Downstream analysis of DNA-seq
6. Understanding publications of DNA-seq
7. Understanding genetic counseling
8. Predict genomics cancer biomarkers and drug targets

COSMIC for clinical NGS testing



<https://digitalinsights.qiagen.com/somatic-mutation-databases-for-clinical/>

COSMIC in the cancer drug discovery



COSMIC for biopharma research



View applications of COSMIC for biopharma research

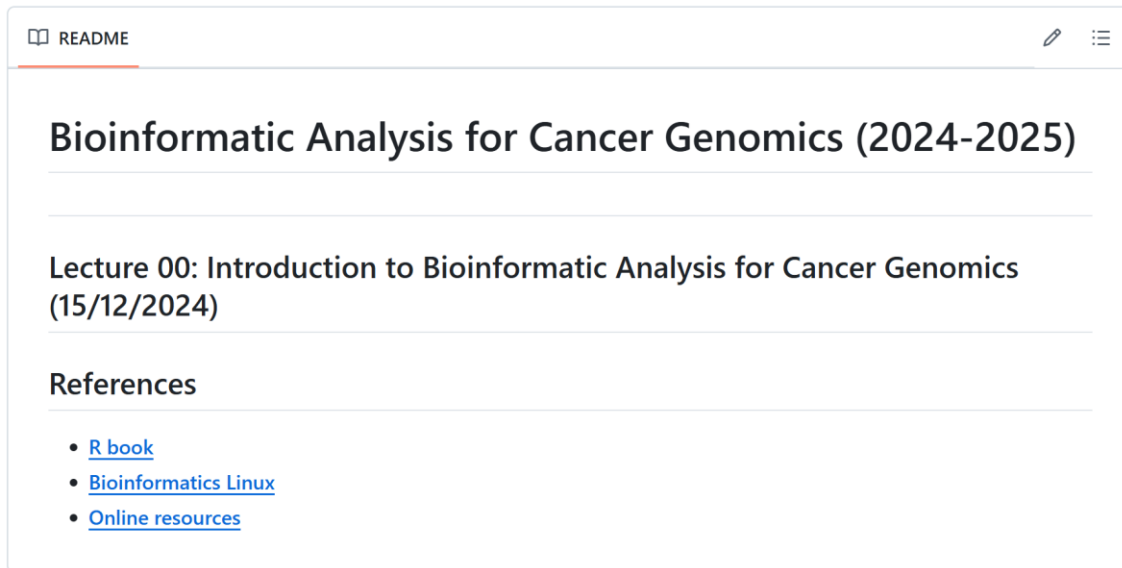
Developed and maintained by Wellcome Sanger Institute, the latest release, [COSMIC v101](#) (November 2024), provides comprehensive coverage of all human genes, featuring over 24.9 million genetic variants derived from 1.5 million tumor samples spanning more than 1,400 cancer types. Notably, the release focused on curating somatic mutation profiles in rare cancers and cancers with significant medical need, as well as the prioritization of studies focusing on whole genome or whole exome data.

- Integrates somatic data from multiple sources, including the International Cancer Genome Consortium, The Cancer Genome Atlas, gene- and genome-wide screens, clinical trials and therapies.
- Data is translated into a standardized format and available through downloadable datasets and user-friendly data exploration tools.
- Every six months, COSMIC content and features are updated to ensure you remain informed on the latest findings.

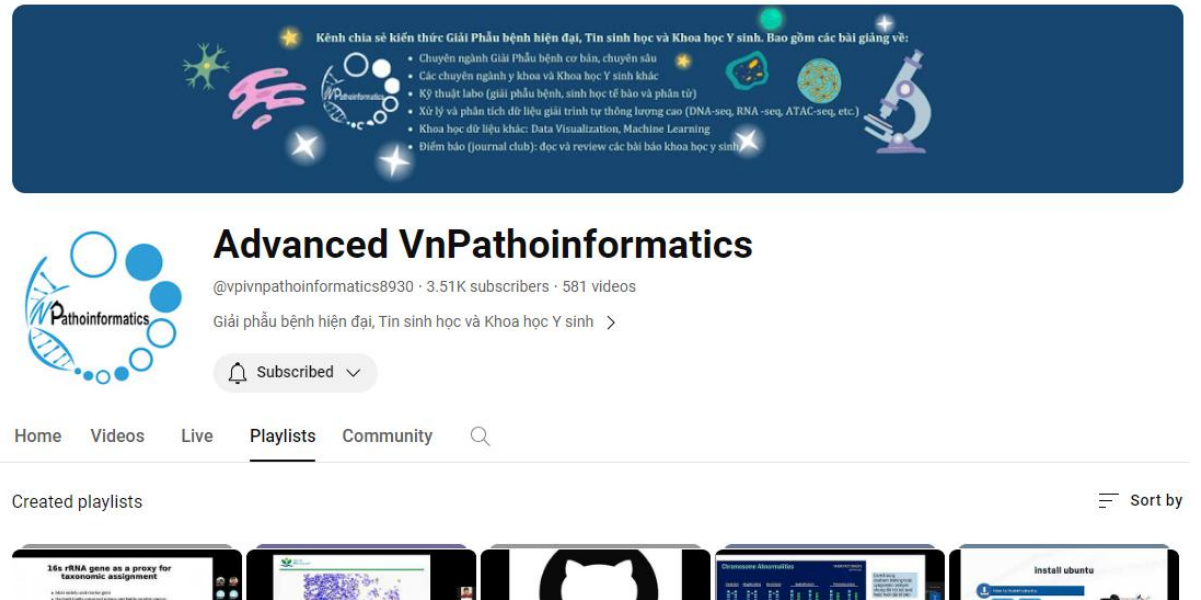
Github and youtube

1. Github: Lecture Slides, Codes/scripts, Exercises, Books and Q&A

https://github.com/luuloi/Bioinformatic_Analysis_for_Cancer_Genomics



The screenshot shows a GitHub repository page. At the top, there is a 'README' tab. The main heading is 'Bioinformatic Analysis for Cancer Genomics (2024-2025)'. Below this, there is a section for 'Lecture 00: Introduction to Bioinformatic Analysis for Cancer Genomics (15/12/2024)'. Underneath, there is a 'References' section with three links: 'R book', 'Bioinformatics Linux', and 'Online resources'.



The screenshot shows a YouTube channel page for 'Advanced VnPathoinformatics'. The channel banner features a blue background with various scientific icons and text in Vietnamese. The channel name is 'Advanced VnPathoinformatics' with a subscriber count of 3.51K and 581 videos. Below the channel name, there is a 'Subscribed' button. The page shows a list of created playlists, including '16s rRNA gene as a proxy for taxonomic assignment', 'Install ubuntu', and others.

2. Youtube:

<https://www.youtube.com/@vpivnpathoinformatics8930/playlists>

Thank you for your listening!

Phuc Loi Luu, PhD

Email: Luu.p.loi@googlemail.com

Zalo: 0901802182